GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 31, 2002, 14:47:54 ; Search time 34.37 Seconds
(without alignments)
2136.897 Million cell updates/sec

Title: US-09-759-056-2
Perfect score: 3478
Sequence: 1 M[illegible]...........[illegible] 667

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111973796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_[illegible]2002:*
1: [illegible],AA1[illegible].DAT:*
2: [illegible],AA1981.DAT:*
3: [illegible],AA1982.DAT:*
4: [illegible],AA1983.DAT:*
5: [illegible],AA1984.DAT:*
6: [illegible],AA1985.DAT:*
7: [illegible],AA1986.DAT:*
8: [illegible],AA1987.DAT:*
9: [illegible],AA1988.DAT:*
10: [illegible],AA1989.DAT:*
11: [illegible],AA1990.DAT:*
12: [illegible],AA1991.DAT:*
13: [illegible],AA1992.DAT:*
14: [illegible],AA1993.DAT:*
15: [illegible],AA1994.DAT:*
16: [illegible],AA1995.DAT:*
17: [illegible],AA1996.DAT:*
18: [illegible],AA1997.DAT:*
19: [illegible],AA1998.DAT:*
20: [illegible],AA1999.DAT:*
21: [illegible],AA2000.DAT:*
22: [illegible],AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3478	100.0	667	22	AAU04557	Human Straб homolo
2	3478	100.0	667	22	AAB88572	Human hydrophobic
3	3412.5	98.1	658	22	AAU04558	Human Straб homolo
4	2903	83.5	560	22	AAB94108	Human protein sequ
5	1271	36.5	247	22	AAM23581	Murine EST encoded
6	860	24.7	178	21	AAB53256	Human colon cancer
7	860	24.7	178	22	AAG74057	Human colon cancer
8	441	12.7	88	20	AAW88559	Secreted protein e
9	441	12.7	88	22	AAB50326	Human secreted pro
10	115.5	3.3	416	22	AAU49235	Propionibacterium
11	114.5	3.3	1067	22	AAB57887	Drosophila melanog
12	114	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
13	114	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
14	112.5	[illegible]	[illegible]	22	[illegible]	[illegible]
15	111.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
16	111	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
17	110.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
18	110	[illegible]	[illegible]	[illegible]	AAM9[illegible]	[illegible]
19	109.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
20	109.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
21	108.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
22	107.5	[illegible]	[illegible]	[illegible]	AAU49694	[illegible]
23	104	[illegible]	[illegible]	22	[illegible]	[illegible]
24	103.5	[illegible]	[illegible]	[illegible]	AAM7[illegible]	[illegible]
25	[illegible]	[illegible]	[illegible]	[illegible]	AAM79004	[illegible]
26	102	[illegible]	[illegible]	[illegible]	AAY32141	[illegible]
27	102	[illegible]	[illegible]	[illegible]	AAB4[illegible]	[illegible]
28	102	[illegible]	[illegible]	[illegible]	AAY[illegible]	[illegible]
29	102	[illegible]	[illegible]	[illegible]	AAM7[illegible]	[illegible]
30	102	[illegible]	[illegible]	[illegible]	AAM93548	[illegible]
31	102	[illegible]	[illegible]	[illegible]	AAY32142	[illegible]
32	102	[illegible]	[illegible]	22	ABB[illegible]	[illegible]
33	102	[illegible]	[illegible]	[illegible]	AAW[illegible]	[illegible]
34	102	[illegible]	[illegible]	[illegible]	AAB[illegible]	[illegible]
35	102	[illegible]	[illegible]	[illegible]	AAW90274	[illegible]
36	101.5	[illegible]	[illegible]	[illegible]	AAY[illegible]	[illegible]
37	101.5	[illegible]	[illegible]	[illegible]	AAB4[illegible]	[illegible]
38	101.5	[illegible]	[illegible]	[illegible]	ABG[illegible]44	[illegible]
39	101.5	[illegible]	[illegible]	22	AAM7890[illegible]	[illegible]
40	101.5	[illegible]	[illegible]	22	AAM78[illegible]	[illegible]
41	100.5	[illegible]	[illegible]	[illegible]	AAY[illegible]	[illegible]
42	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
43	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]2804	[illegible]
44	[illegible]	[illegible]	[illegible]	[illegible]	AAB[illegible]	[illegible]
45	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

ALIGNMENTS

RESULT 1
AAU04557
ID AAU04557 standard; [illegible] AA.
XX
AC AAU04557;
XX
DT 26-SEP-2001 (first entry)
XX
DE Human Stra6 [illegible]
XX
KW Human; Stra6; [illegible] tumour; [illegible]
KW head tumour; [illegible]
KW immunogenic [illegible]
XX
OS Homo sapiens.
XX
FH Key [illegible]
FT Modified site [illegible]
FT /note= "Asn is N-glycosylat[illegible]"
FT Region [illegible]
FT /label= N-glycosylat[illegible]
FT Domain [illegible]
FT /label= transmembrane [illegible]
FT Region [illegible]
FT /note= "[illegible] deleted [illegible]"
FT Domain [illegible]
FT /label= transmembrane doma[illegible]
FT Region [illegible]
FT /note= "[illegible]"
FT Domain [illegible]
FT /label= transmembrane [illegible]
FT Region [illegible]
FT /label= N-glycosylat[illegible]
FT Domain [illegible]

```
FT                   /label- Transmembrane_domain
FT   Region          241..247
FT                   /label- N_myristoylation_site
FT   Domain          301..320
FT                   /label- Transmembrane_domain
FT   Region          317..323
FT                   /label- N_myristoylation_site
FT   Region          341..347
FT                   /label- N_myristoylation_site
FT   Region          364..375
FT                   /note= "Prokaryotic membrane lipoprotein attachment
FT                   site"
FT   Domain          364..380
FT                   /label- Transmembrane_domain
FT   Domain          431..451
FT                   /label- Transmembrane_domain
FT   Domain          474..489
FT                   /label- Transmembrane_domain
FT   Region          525..531
FT                   /label- N_myristoylation_site
FT   Domain          568..585
FT                   /label- Transmembrane_domain
FT   Region          627..633
FT                   /label- N_myristoylation_site
FT   Region          631..637
FT                   /label- N_myristoylation_site
FT   Region          640..646
FT                   /label- N_myristoylation_site
FT   Region          661..667
FT                   /label- N_myristoylation_site
XX
PN   WO200151635-A2.
XX
PD   19-JUL-2001.
XX
PF   11-JAN-2001; 2001WO-US00847.
XX
PR   13-JAN-2000; 2000US-0175849.
PR   14-APR-2000; 2000US-0197089.
PR   29-AUG-2000; 2000US-0228914.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Pennica D,  Smith V,  Wood WI;
XX
DR   WPI; 2001-442146/47.
DR   N-PSDB; AAS08640.
XX
PT   Human nucleic acid encoding a PRO10282 or PRO19578 polypeptide (the PRO
PT   polypeptides are referred to as Stra6 polypeptides), useful in
PT   molecular biology, including use as hybridisation probes, and in
PT   chromosome and gene mapping -
XX
PS   Claim 24; Fig 2; 159pp; English.
XX
CC   The sequence is human PRO10282, a homologue of murine Stra6, a
CC   retinoic acid responsive protein thought to play an important role in
CC   early dorsoventral limb patterning during development and later in the
CC   control of endochondral ossification. The gene for the Stra6 homologue
CC   is located on chromosome 15q23. The Stra6 polypeptides, agonists,
CC   antagonists or anti-Stra6 antibodies are useful for preparation of a
CC   medicament useful in the treatment of a condition which is responsive to
CC   the Stra6 polypeptides, agonists, antagonists or anti-Stra6  antibodies.
CC   The Stra6 polypeptides may also be employed as molecular weight markers
CC   for protein electrophoresis. The Stra6 nucleic acids have applications in
CC   molecular biology, including use as hybridisation probes, and in
CC   chromosome and gene mapping. The antibodies and other anti-tumour
CC   compounds may be used to treat various conditions, including those
CC   characterised by overexpression and/or activation of amplified genes.
CC   Exemplary conditions or disorders include benign or malignant tumours
CC   (e.g., renal, liver, kidney, bladder, breast, gastric, ovarian,
CC   colorectal, prostate, pancreatic, lung, vulval, thyroid, hepatic
CC   carcinomas, sarcomas, glioblastomas, and various head and neck tumours,
CC   leukaemias and lymphoid malignancies), neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal and
CC   blastocoelic disorders, inflammatory, angiogenic and immunogenic
CC   disorders.
XX
SQ   Sequence   667 AA;

  Query Match             100.0%;  Score 3478;  DB 22;  Length 667;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  667;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSSQPAGNQTSPGATEDYSYGSWYIDEPQGGEELQPEGEVPSCHTSIPPGLYHACLASLS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 mssqpagnqtspgatedysygswyidepqggeelqpegevpschtsippglyhaclasls 60

Qy         61 ILVLLLLAMLVRRRQLWPDCVRGRPGLPSPVDFLAGDRPRAVPAAVFMVLLSSLCLLLPD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ilvllllamlvrrrqlwpdcvrgrpglpspvdflagdrpravpaavfmvllsslclllpd 120

Qy        121 EDALPFLTLASAPSQDGKTEAPRGAWKILGLFYYAALYYPLAACATAGHTAAHLLGSTLS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 edalpfltlasapsqdgkteaprgawkilglfyyaalyyplaacataghtaahllgstls 180

Qy        181 WAHLGVQVWQRAECPQVPKIYKYYSLLASLPLLLGLGFLSLWYPVQLVRSFSRRTGAGSK 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 wahlgvqvwqraecpqvpkiykyysllaslplllglgflslwypvqlvrsfsrrtgagsk 240

Qy        241 GLQSSYSEEYLRNLLCRKKLGSSYHTSKHGFLSWARVCLRHCIYTPQPGFHLPLKLVLSA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 glqssyseeylrnllcrkklgssyhtskhgflswarvclrhciytpqpgfhlplklvlsa 300

Qy        301 TLTGTAIYQVALLLLVGVVPTIQKVRAGVTTDVSYLLAGFGIVLSEDKQEVVELVKHHLW 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 tltgtaiyqvallllvgvvptiqkvragvttdvsyllagfgivlsedkqevvelvkhhlw 360

Qy        361 ALEVCYISALVLSCLLTFLVLMRSLVTHRTNLRALHRGAALDLSPLHRSPHPSRQAIFCW 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 alevcyisalvlsclltflvlmrslvthrtnlralhrgaaldlsplhrsphpsrqaifcw 420

Qy        421 MSFSAYQTAFICLGLLVQQIIFFLGTTALAFLVLMPVLHGRNLLLFRSLESSWPFWLTLA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 msfsayqtaficlgllvqqiifflgttalaflvlmpvlhgrnlllfrslesswpfwltla 480

Qy        481 LAVILQNMAAHWVFLETHDGHPQLTNRRVLYAATFLLFPLNVLVGAMVATWRVLLSALYN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 lavilqnmaahwvflethdghpqltnrrvlyaatfllfplnvlvgamvatwrvllsalyn 540

Qy        541 AIHLGQMDLSLLPPRAATLDPGYYTYRNFLKIEVSQSHPAMTAFCSLLLQAQSLLPRTMA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 aihlgqmdlsllppraatldpgyytyrnflkievsqshpamtafcslllqaqsllprtma 600

Qy        601 APQDSLRPGEEDEGMQLLQTKDSMAKGARPGASRGRARWGLAYTLLHNPTLQVFRKTALL 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 apqdslrpgeedegmqllqtkdsmakgarpgasrgrarwglaytllhnptlqvfrktall 660

Qy        661 GANGAQP 667
              |||||||
Db        661 ganqaqp 667


RESULT 2
AAB88572
ID   AAB88572 standard; Protein; 667 AA.
XX
AC   AAB88572;
XX
DT   04-JUN-2001  (first entry)
XX
DE   Human hydrophobic domain containing protein clone HP10713 #36.
XX
KW   Human; hydrophobic domain; immunosuppressant; anti-HIV; neuroprotective;
KW   antianaemic; vulnerary; antiulcer; osteopathic; anti-inflammatory;
```

```
KW   cytostatic; gene therapy; autoimmune disorder; multiple sclerosis;
KW   HIV infection, anaemia, burn, ulcer, osteoporosis  tumour; wound healing;
KW   inflammatory bowel disease; nutritional supplement; appetite; vaccine;
KW   behavioural characteristic; immune response.
XX
OS   Homo sapiens.
XX
PN   WO200112[illegible]-A2.
XX
PD   22-FEB-2001.
XX
PF   1[illegible]-AUG-2000; 2000WO-JP05[illegible].
XX
PR   17-AUG-1999;   99JP-0230344.
PR   07-SEP-1999;   99JP-0252551.
PR   01-OCT-1999;   99JP-0281132.
PR   22-OCT-1999;   99JP-0301624.
PR   04-NOV-1999;   99JP-0313877.
XX
PA   (SAGA ) SAGAMI CHEM RES CENT.
PA   (PROT-) PROTEGENE INC.
XX
PI   Kato S,  Kimura T;
XX
DR   WPI; 2001-160059/16.
DR   N-PSDB; AAF94442.
XX
PT   Human proteins with hydrophobic domains and the DNAs which encode them
PT   are useful for treating autoimmune disorders, burns and tumors and for
PT   screening novel pharmaceuticals
XX
PS   Claim 1; Page 288-291; 518pp; English.
XX
CC   AAF94417 to AAF94516 encode the human proteins given in AAB88557 to
CC   AAB886[illegible] (I) which have a hydrophobic domain. (I) have immunosuppressant,
CC   anti-HIV, neuroprotective, antianaemic, vulnerary, antiulcer,
CC   osteopathic, anti-inflammatory and cytostatic activities, and can be
CC   used in gene therapy. (I) can be used as pharmaceuticals and as antigens
CC   to prepare antibodies. DNA and cDNA (II) encoding (I) can be used as
CC   probes for genetic diagnosis and gene sources for gene therapy or for
CC   producing (I) in large quantities. Cells containing (II) are used for
CC   the detection of ligands or receptors corresponding to membrane or
CC   secretory proteins and to screen small molecule novel pharmaceuticals.
CC   Antibodies directed to (I) can be used for the detection, quantification
CC   and purification of (I). Activities of (I) may include cytokine and cell
CC   proliferation/differentiation function, immune stimulating or suppressing
CC   activity, haematopoiesis regulating activity, tissue growth activity,
CC   activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, receptor/ligand activity and anti-inflammatory
CC   activity. (I) and (II) can be used to treat autoimmune disorders e.g.
CC   multiple sclerosis, HIV infections, anaemia, burns, ulcers, osteoporosis,
CC   inflammatory bowel disease and tumours. (I) and (II) can also be used for
CC   wound healing, as nutritional sources or supplements e.g. as amino acid,
CC   carbon or nitrogen source, to effect metabolism, catabolism, anabolism,
CC   processing and utilisation of dietary fat, protein, carbohydrate,
CC   vitamins and minerals, to effect behavioural characteristics, to affect
CC   appetite, and can act as antigens in vaccines to raise an immune response
CC   to the protein or another material cross reactive with the protein
XX
SQ   Sequence   667 AA.

  Query Match              100.0%;  Score 3478;  DB 22;  Length 667;
  Best Local Similarity    100.0%;  Pred. No. 0;
  Matches  667;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MSSQPAGNQTSPGATEDYSYGSWYIDEPQGGEELQPEGEVPSCHTSIPPGLYHACLASLS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 mssqpagnqtspgatedysyqswyidepqgqeelqpeqevpschtsippglyhaclasls 60

Qy         61 ILVLLLLAMLVRRRQLWPDCVRGRPGLPSPVDFLAGDRPRAVPAAVFMVLLSSLCLLLPD 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ilvlllamlvrrrqlwpdcvrgrpglpspvdflagdrpravpaavfmvllsslclllpd 120

Qy        121 EDALPFL[illegible]
Db        121 edalpfl[illegible]

Qy        181 WAHLGVQVW[illegible]
Db        181 wahlgv[illegible]

Qy        241 GLQSSYS[illegible]
Db        241 glqssys[illegible]

Qy        301 TLTGTALY[illegible]
Db        301 tltgtaly[illegible]

Qy        361 ALEVCYISA[illegible]
Db        361 alevcyisa[illegible]

Qy        421 MSFSAYQ[illegible]
Db        421 msfsayq[illegible]

Qy        481 LAVILLQNMAA[illegible]
Db        481 lavillqnma[illegible]

Qy        541 AIHLGQM[illegible]
Db        541 aihlqqm[illegible]

Qy        601 APQDSLP[illegible]
Db        601 apqdslp[illegible]

Qy        661 GANGAQP[illegible]
Db        661 gangaqp[illegible]


RESULT 3
AAU04558
ID   AAU04558 standard; [illegible] 658 AA.
XX
AC   AAU04558;
XX
DT   26-SEP-2001  (first entry)
XX
DE   Human Straf[illegible]
XX
KW   Human; Straf[illegible]
KW   head tumour; [illegible]
KW   immunogenic [illegible]; antigen; antib[illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified site   [illegible]
FT                   [illegible] N [illegible]
FT   Region          [illegible]
FT                   /label= N-myristoylation [illegible]
FT   Domain          [illegible]
FT                   /label= transmembrane [illegible]
FT   Domain          [illegible]
FT                   /label= transmembrane [illegible]
FT   Region          [illegible]
FT                   [illegible]
FT   Domain          [illegible]
FT                   /label= transmembrane [illegible]
FT   Region          [illegible]
FT                   /label= N-myristoylation [illegible]
```

```
FT   Domain          197..214
FT                   /label= Transmembrane_domain
FT   Region          232..238
FT                   /label= N_myristoylation_site
FT   Domain          291..312
FT                   /label= Transmembrane_domain
FT   Region          308..314
FT                   /label= N_myristoylation site
FT   Region          332..338
FT                   /label= N_myristoylation site
FT   Region          355..366
FT                   /note= "Prokaryotic membrane lipoprotein attachment
FT                   site"
FT   Domain          356..371
FT                   /label= Transmembrane_domain
FT   Domain          425..444
FT                   /label= Transmembrane_domain
FT   Domain          464..481
FT                   /label= Transmembrane_domain
FT   Domain          505..522
FT                   /label= Transmembrane_domain
FT   Region          516..522
FT                   /label= N_myristoylation_site
FT   Region          618..624
FT                   /label= N_myristoylation site
FT   Region          622..628
FT                   /label= N_myristoylation_site
FT   Region          631..637
FT                   /label= N_myristoylation_site
FT   Region          652..658
FT                   /label= N_myristoylation_site
XX
PN   WO200151635 A2.
XX
PD   19-JUL-2001.
XX
PF   11-JAN-2001; 2001WO-US00847.
XX
PR   13-JAN-2000; 2000US-0175644.
PR   14-APR-2000; 2000US-0197089.
PR   24-AUG-2000; 2000US-0228914.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Fennica D,  Smith V,  Wood WI;
XX
DR   WPI; 2001-442146/47.
DR   N-PSDB; AAS08632.
XX
PT   Human nucleic acid encoding a PRO10282 or PRO19578 polypeptide (the PRO
PT   polypeptides are referred to as Stra6 polypeptides), useful in
PT   molecular biology, including use as hybridisation probes, and in
PT   chromosome and gene mapping
XX
PS   Claim 24; Fig 7; 159pp; English.
XX
CC   The sequence is human PRO19578 (a variant possibly due to
CC   alternative splicing of the Stra6 mRNA), a homologue of murine Stra6, a
CC   retinoic acid responsive protein thought to play an important role in
CC   early dorsoventral limb patterning during development and later in the
CC   control of endochondral ossification. The gene for the Stra6 homologue
CC   is located on chromosome 15q23. The Stra6 polypeptides, agonists,
CC   antagonists or anti-Stra6 antibodies are useful for preparation of a
CC   medicament useful in the treatment of a condition which is responsive to
CC   the Stra6 polypeptides, agonists, antagonists or anti Stra6 antibodies.
CC   The Stra6 polypeptides may also be employed as molecular weight markers
CC   for protein electrophoresis. The Stra6 nucleic acids have applications in
CC   molecular biology, including use as hybridisation probes, and in
CC   chromosome and gene mapping. The antibodies and other anti-tumour
CC   compounds may be used to treat various conditions, including those
CC   [illegible]
CC   Exemplary conditions or disorders include benign or malignant tumours
CC   (e.g., renal, liver, kidney, bladder, breast, gastric, ovarian,
CC   colorectal, prostate, pancreatic, lung, vulval, thyroid, hepatic
CC   carcinomas, sarcomas, glioblastomas, and various head and neck tumours,
CC   leukaemias and lymphoid malignancies), neuronal, glial, astrocytal,
CC   hypothalamic and other glandular, macrophagal, epithelial, stromal and
CC   blastocoelic disorders, inflammatory, angiogenic and immunogenic
CC   disorders.
XX
SQ   Sequence   658 AA;

  Query Match             98.1%;  Score 3412.5;  DB 22;  Length 658;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches  657;  Conservative    1;  Mismatches    0;  Indels    9;  Gaps    1;

Qy          1 MSSQPAGNQTSPGATEDYSYGSWYIDEPQGGEELQPEGEVPSCHTSIPPGLYHACLASLS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 mssqpagnqtspgatedysygswyideppqgeelqpeqevpschtsippglyhaclasls 60

Qy         61 ILVLLLLAMLVRRRQLWPDCVRGRPGLPS[illegible] 120
              ||||||||||||||||||||||||||||         ||||||||||||||||||||||
Db         61 ilvllllamlvrrrqlwpdcvrgrpglp---------[illegible]ipd 111

Qy        121 [illegible] 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 [illegible] 171

Qy        181 [illegible] 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 [illegible] 231

Qy        241 [illegible] 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 [illegible] 291

Qy        301 [illegible] 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 [illegible] 351

Qy        361 [illegible] 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 alevcyisalvlsclltt[illegible] 411

Qy        421 [illegible] 480
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        412 mstsayqtaf[illegible]sswp[illegible] 471

Qy        481 [illegible] 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 [illegible] 531

Qy        541 [illegible] 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 [illegible] 591

Qy        601 [illegible] 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        592 [illegible] 651

Qy        661 CANQAQP 667
              |||||||
Db        652 canqaqp 658


RESULT 4
AAB94108
ID   AAB94108 standard; Protein; 560 AA.
XX
AC   AAB94108;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human protein sequence SEQ ID NO:14340.
XX
```

```
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 14340; 2537pp + CD-ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence   560 AA;

  Query Match             83.5%;  Score 2503;  DB 22;  Length 560;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-297;
  Matches  560;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        108 MVLLSSLCLLLPDPDALPPLTLASAPSQGKTEAPPGAWKILGLFYYAALYYPLAACATA 167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 mvllsslclllpdpdalppltlasapsqgkteappgawkilglfyyaalyyplaacata 60

Qy        168 GHTAAHLLGSTLSWAHLGVQVWQRAECPQVPKIYKYYSLLASLPLLLGLGFLSLWYPVQL 227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ghtaahllgstlswahlgvqvwqraecpqvpkiykyysllaslplllglgflslwypvql 120

Qy        228 VRSFSRRTGAGSKGLQSSYSGEEYLRNLLCRKKLGSSYHTSKHGFLSWARVCLRHCIYTPQ 287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 vrsfsrrtgagskglqssysgeeylrnllcrkklgssyhtskhgflswarvclrhciytpq 180

Qy        288 PGFHLPLKLVLSALLLGIAIYQVALLLLVGVVPTLQKVRAGVTTDVSYLLAGFGIVLSED 347
              [illegible]
Db        181 pgfhlplk[illegible]

Qy        348 KLEVVELVE[illegible]
Db        241 kqevvelve[illegible]

Qy        408 RSPHPSR[illegible]
Db        301 rsphpsr[illegible]

Qy        468 SLESSWPTW[illegible]
Db        361 slesswptw[illegible]

Qy        528 VATWRVLL[illegible]
Db        421 vatwrvll[illegible]

Qy        588 LLQAQSL[illegible]
Db        481 llqaqsl[illegible]

Qy        648 NPTLQVEFRTA[illegible]
Db        541 nptlqvtrktaflaa[illegible]

RESULT 5
AAM23581
ID   AAM23581 standard; protein; 347 AA.
XX
AC   AAM23581;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Murine EST encoded protein SEQ ID NO: [illegible]
XX
KW   Human; sheep; [illegible] fruit fly; yeast; [illegible]
KW   tomato; monkey; [illegible] expressed [illegible]
KW   diagnostics; [illegible] gene mapping; [illegible]
KW   biodiversity; gene therapy; nutrition.
XX
OS   Mus musculus.
XX
PN   WO200154477-A.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; [illegible]
XX
PR   21-JAN-2000; [illegible]
PR   17-JUL-2000; [illegible]
PR   03-AUG-2000; [illegible]
PR   15-SEP-2000; [illegible]
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, [illegible] Wang [illegible]
PI   Cao Y, Drmanac RT [illegible] Wehrman [illegible]
XX
DR   WPI; 2001-47[illegible]
DR   N-PSDB; AAH[illegible]
XX
PT   Isolated polypeptide [illegible] treatment [illegible]
PT   antibodies and research [illegible]
XX
PS   Claim 20; [illegible] English.
XX
CC   The present invention provides the proteins [illegible]
CC   proteins [illegible]
CC   cow, pig, hamster [illegible] yeast [illegible]
CC   urchin and [illegible] were derived [illegible]
```

CC from the organism of interest. They can be used in diagnostics, forensics, gene mapping, identification of mutations, to assess biodiversity and for nutritional purposes. The present sequence is a protein of the invention.
XX
SQ Sequence 247 AA;

Query Match 86.5%; Score 1271; DB 22; Length 247;
Best Local Similarity 100.0%; Pred. No. 1.5e-125;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 421 MSFSAYQTAFICIGLLVQQTIFFLGFTIALAFL 455
Db 1 matsayqtaficlqllvqqiiffigftialafl 35

Qy 456 VLMPVLHEGRNLLFRGLEGWFPWLCLALAVILQNMAABWVFLETHLGHPGLCNRFVLYA 515
Db 36 vlmpvlhegrnllfrglegwfpwlclalavilqnmaabwvfletHlghpglcnrfvlya 95

Qy 516 AIFLLFFLNVLVGAMVAIWEVLLSALYNAIHEQKESAFVKEAIDP 561
Db 96 aiflllffLnvlvgamvaiwevllsalynaihlqqmdlsllppraalstp 141

XX
AC AAG97991 to AAG98763 encode the human colon cancer associated proteins, called human colon cancer antigens, given in AAB54004 to AAB54026. The human colon cancer antigens can have cytostatic, cardioactive, muscular, neuroprotective, immunomodulatory, gynaecological, gastrointestinal, vulnerary, nephrotropic, antiinfective and antibacterial activities, and can be used in gene therapy. The colon cancer antigen polynucleotides, proteins and antibodies to the proteins are useful for the prevention, treatment and diagnosis of colon disorders, such as colon cancer. The polynucleotides may be used in diagnostics and research, such as for chromosome identification, and as hybridisation probes. The proteins may also be used to prevent diseases such as neural disorders, immune system disorders, muscular disorders, reproductive disorders, gastrointestinal disorders, wounds, renal disorders, infectious diseases, and cardiovascular disorders. AAG98764 to AAG98772 and AAB54007 represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 178 AA;

Query Match 34.2%; Score 863; DB 21; Length 178;
Best Local Similarity 99.2%; Pred. No. 2.5e-82;
Matches 166; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 393 RALHRGAALDLSPLHRSPHPSRQAIPCWMSFSAYQTAFICIGLLVQQTIFFLGFTIALAFL 452
Db 2 ralhrqaaldlsplhrsphpsrqaitcwmstsayqtaficlqllvqqiiffigftialafl 61

Qy 453 VLMPVLHEGRNLLFRGLEGWFPWLCLALAVILQNMAABWVFLETHLGHPGLCNRFVLYA 512
Db 62 vlmpvlhegrnllfrglegwfpwlclalavilqnmaabwvfletHlghpglcnrfvlya 121

Qy 513 AIFLLFFLNVLVGAMVAIWEVLLSALYNAIHEQKESAFVKEAIDP 558
Db 122 aiflllfplnvlvgamvaiwevllsalynaihlqqmdlsllppraalstp 167

RESULT 7
AAG74057
ID AAG74057 standard; Protein; 178 AA.
XX
AC AAG74057;
XX
DT 09-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:4821.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection; colorectal carcinoma.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US026524
XX
PR 29-SEP-1999; 99US-0407737.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-254137/24.
DR N-PSDB; AAB33488.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides, useful for preventing, diagnosing and/or treating colorectal cancers
XX
PS Claim 11; Page 6606-6607; 9803pp; English.

```
XX
CC   AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC   cancer associated nucleic acid molecules (N) and proteins (P), where
CC   the proteins are collectively known as colon cancer antigens. The colon
CC   cancer antigens have cytostatic activity and can be used in gene
CC   therapy and vaccine production. N and P may be used in the prevention,
CC   diagnosis and treatment of diseases associated with inappropriate P
CC   expression. For example, N and P may be used to treat disorders
CC   associated with decreased expression by rectifying mutations or deletions
CC   in a patient's genome that affect the activity of P by expressing
CC   inactive proteins or to supplement the patients own production of P.
CC   Additionally, N may be used to produce the colon cancer-associated Ps,
CC   by inserting the nucleic acids into a host cell and culturing the cell
CC   to express the proteins. N and P can be used in the prevention, diagnosis
CC   and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC   and AAG77789 represent sequences used in the exemplification of the
CC   present invention.
CC   N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC   missing at time of publication, meaning no sequences are present for
CC   SEQ ID No:1627 to 1652, 7921 and 7922.
XX
SQ   Sequence   178 AA;

  Query Match             24.7%;  Score 860;  DB 22;  Length 178;
  Best Local Similarity   98.2%;  Pred. No. 2.5e-82;
  Matches  166;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        393 RALHRQAAIDISPIHRSPHPSRQAIEWMSTSAYQTATIETQLLVQQIIPPIQTTALATL 452
Db          2 ralhrqaaidispihrsphpsrqaitewmstsayqtatietqllvqqiittlqttalatl 61

Qy        453 VLMPVLHGRNLLLFRSLESSWPFWLILALAVILQNMAAHWVFLETHDGHPQLINRRVLYA 512
Db         62 vlmpvlharnllltrslessvptwltlalavilqnmaahwvfletndghpqltnrrvlya 121

Qy        513 AIPLLFPYNVLVGAMVAINRVLLSALYNAIHLQQMDLSLLPRAAILDP 561
Db        122 atllitplnvlvqamvatwrvllsalynaihlqqmdlsllppraalstp 170


RESULT    8
AAW88559
ID   AAW88559 standard; Protein; 88 AA.
XX
AC   AAW88559;
XX
DT   01-MAR-1999  (first entry)
XX
DE   Secreted protein encoded by gene 26 clone HIDAF28.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS   Homo sapiens.
XX
PN   WO9854963-A2.
XX
PD   10-DEC-1998.
XX
PF   04-JUN-1998;   98WO-US11422.
XX
PR   18-DEC-1997;   97US-0070923.
PR   06-JUN-1997;   97US-0048877.
PR   06-JUN-1997;   97US-0048881.
PR   06-JUN-1997;   97US-0048884.
PR   06-JUN-1997;   97US-0048893.
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   06-JUN-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
PR   05-SEP-1997;   97US-[illegible]
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Brewer LA, [illegible]
PI   Fan P, Feng P, Ferrie AM, Fischer [illegible]
```

```
PI   Florence K, Greene JM, Hu J, Kyaw H, LaFleur DW;
PI   Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM,
PI   Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
XX
DR   WPI; 1999-450865/05.
DR   N-PSDB; AAV84436.
XX
PT   New isolated human genes and the secreted polypeptides they encode
PT   useful for diagnosis and treatment of e.g. cancers, neurological
PT   disorders, immune diseases, inflammation or blood disorders
XX
PS   Claim 11; Page 502; 772pp; English.
XX
CC   The invention relates to nucleic acid sequences (AAV84411 to AAV84633)
CC   encoding human secreted proteins (AAW88534 to AAW88756). The secreted
CC   protein gene sequences are deposited with the ATCC under deposit numbers
CC   ATCC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC   209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
CC   cells comprising recombinant vectors containing the nucleic acid
CC   sequences are used for the recombinant production of the secreted
CC   proteins. The polynucleotide and amino acid sequences are useful for are
CC   useful for preventing, treating or ameliorating medical conditions e.g.
CC   by protein or gene therapy. Pathological conditions can be also
CC   diagnosed by determining the amount of the new polypeptides in a sample
CC   or by determining the presence of mutations in the new polynucleotides.
CC   Specific uses are described for each of the polynucleotides, based on
CC   which tissues they are most highly expressed in, and include developing
CC   products for the diagnosis or treatment of cancer, neurodegenerative
CC   disorders, developmental abnormalities and foetal deficiencies, blood
CC   disorders, tumours, leukemias, diseases of the immune system, autoimmune
CC   diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
CC   ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC   restenosis, prostate diseases, obesity, disorders involving osteoclasts
CC   such as osteoporosis, arthritis or malignancies, diseases of testes,
CC   lung or thymus, digestive/endocrine disorders, infections and AIDS. The
CC   polypeptides are also useful for identifying their binding partners.
CC   The present sequence represents human secreted protein (see descriptor
CC   line for gene number and clone identification).
XX
SQ   Sequence   88 AA;

  Query Match             12.7%;  Score 441;  DB 20;  Length 88;
  Best Local Similarity   98.9%;  Pred. No. 1.7e-38;
  Matches   86;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        581 MTAFCSLLLQAQSLLPRTMAAPQDSLRPGEEDEGMQLLQTKDSMAKGARPGASRGRARWG 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 mtafcslllqaqsllprtmaapqdslrpqeedeqmqllqtkdsmakqarpqaxrqrarwq 60

Qy        641 LAYTLLHNPTLQVFRKTALLGANGAQP 667
              |||||||||||||||||||||||||||
Db         61 laytllhnptlqvfrktallqanqaqp 87


RESULT 9
ABB50326
ID   ABB50326 standard; Protein; 88 AA.
XX
AC   ABB50326;
XX
DT   07-FEB-2002  (first entry)
XX
DE   Human secreted protein encoded by gene 26 SEQ ID NO:274.
XX
KW   Human; secreted protein; immunomodulatory; antisclerotic; anti HIV;
KW   dermatological; immunosuppressive; antiinflammatory; immunostimulant;
KW   cytostatic; cardiant; vasular; anti-angiogenic; ophthalmological;
KW   neuroprotective; nootropic; anticonvulsant; antialzheimers; vulnerary;
KW   antiparkinsonian; antimicrobial; gene therapy; vaccine; immune disorder;
KW   multiple sclerosis; systemic lupus erythematosus; HIV infection; cancer;
KW   human immunodeficiency virus; hyperproliferative disorder; wound healing;
KW   Gaucher's disease; cardiovascular disease; Scimitar syndrome; chemotaxis;
KW   Chaga's cardiomyopathy; coronary arteriosclerosis; angiogenic disorder;
KW   corneal graft neovascularisation; diabetic retinopathy; regeneration;
KW   neurological disorder; Huntington's chorea; Alzheimer's disease;
KW   Parkinson's disease; infectious disease.
XX
OS   Homo sapiens.
XX
PN   WO200162891-A2.
XX
PD   30-AUG-2001.
XX
PF   21-FEB-2001; 2001WO-US05614.
XX
PR   24-FEB-2000; 2000US-184836P.
PR   29-MAR-2000; 2000US-193170P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ni J, Ebner R, LaFleur DW, Moore PA, Olsen HS, Rosen CA;
PI   Ruben SM, Soppet DR, Young PE, Shi Y, Florence KA, Wei Y;
PI   Florence C, Hu J, Li Y, Kyaw H, Fischer CL, Ferrie AM, Fan P;
PI   Feng P, Endress GA, Dillon PJ, Carter KC, Brewer LA, Yu G;
PI   Zeng Z, Greene JM;
XX
DR   WPI; 2001-625724/72.
DR   N-PSDB; ABA83219.
XX
PT   Nucleic acids encoding 297 human secreted polypeptides, useful for
PT   preventing, diagnosing and/or treating, e.g. cancers, Parkinson's
PT   disease and diabetic retinopathy -
XX
PS   Claim 11; Page 1092; 1533pp; English.
XX
CC   ABB50401 to ABB51287 and ABA83194 to ABA83441 represent human secreted
CC   proteins (I) and polynucleotide (II) sequences. (I) and (II) have various
CC   activities based on the tissues and cells the genes are expressed in.
CC   Example of these activities include: immunomodulatory; antisclerotic;
CC   dermatological; immunosuppressive; antiinflammatory; immunostimulant;
CC   anti-HIV; cytostatic; cardiant; anti angiogenic; ophthalmological;
CC   neuroprotective; nootropic; anticonvulsant; antialzheimers; vascular;
CC   antiparkinsonian; antimicrobial; and vulnerary. (I) and (II) can be used
CC   in gene therapy and vaccine production. (I) and (II) can be used in the
CC   prevention, diagnosis and treatment of immune disorders (e.g. multiple
CC   sclerosis, systemic lupus erythematosis and human immunodeficiency virus
CC   (HIV) infections), hyperproliferative disorders (e.g. cancers and
CC   Gaucher's disease), cardiovascular diseases (e.g. Scimitar syndrome,
CC   Chaga's cardiomyopathy, and coronary arteriosclerosis), angiogenic
CC   disorders (e.g. corneal graft neovascularisation and diabetic
CC   retinopathy), neurological disorders (e.g. Huntington's chorea,
CC   Alzheimer's disease and Parkinson's disease), infectious diseases and/or
CC   for promoting wound healing, regeneration and/or chemotaxis. ABA83185 to
CC   ABA83193 and ABB50400 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence   88 AA;

  Query Match             12.7%;  Score 441;  DB 22;  Length 88;
  Best Local Similarity   98.9%;  Pred. No. 1.7e-38;
  Matches   86;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        581 MTAFCSLLLQAQSLLPRTMAAPQDSLRPGEEDEGMQLLQTKDSMAKGARPGASRGRARWG 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 mtafcslllqaqsllprtmaapqdslrpqeedeqmqllqtkdsmakqarpqaxrqrarwq 60

Qy        641 LAYTLLHNPTLQVFRKTALLGANGAQP 667
              |||||||||||||||||||||||||||
Db         61 laytllhnptlqvfrktallqanqaqp 87


RESULT 10
AAU49235
ID   AAU49235 standard; Protein; 416 AA.
```

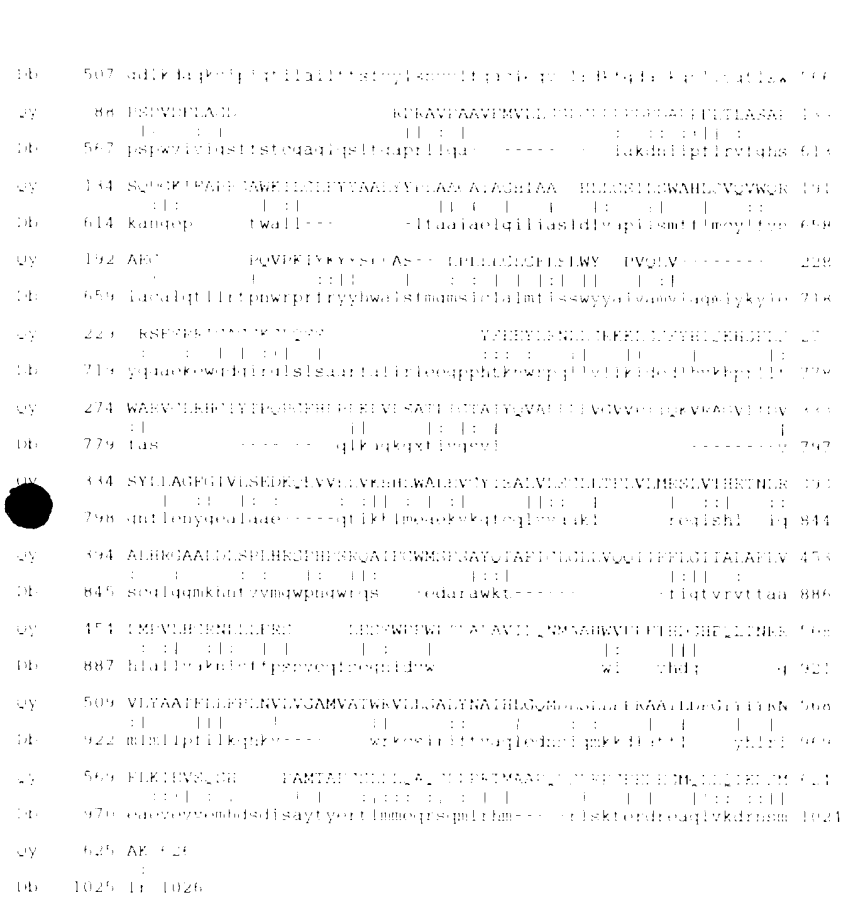
```
XX
AC   AAB4[illegible]
XX
DT   27-FEB-2002  (first entry)
XX
DE   Propionibacterium acnes immunogenic protein #1013.
XX
KW   SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KW   uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW   inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW   dermatological; osteopathic; neuroprotectant.
XX
OS   Propionibacterium acnes.
XX
PN   WO200181581-A2.
XX
PD   01-NOV-2001.
XX
PF   23-APR-2001; 2001WO-US12865.
XX
PR   21-APR-2000; 2000US-199047P.
PR   02-JUN-2000; 2000US-208841P.
PR   07-JUL-2000; 2000US-216747P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Persing DH,  Mitcham JL,  Wang SS,  Bhatia A;
PI   L'maisonneuve J,  Zhang Y,  Jen S,  Carter D;
XX
DR   WPI; 2001-[illegible]774/71.
DR   N-PSDB; AAS59545.
XX
PT   Propionibacterium acnes polypeptides and nucleic acids useful for
PT   vaccinating against and diagnosing infections, especially useful for
PT   treating acne vulgaris
XX
PS   Example 1; SEQ ID NO 10430; 1069pp; English.
XX
CC   Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC   polypeptides. The proteins and their associated DNA sequences are used in
CC   the treatment, prevention and diagnosis of medical conditions caused by
CC   P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC   pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC   P. acnes is also involved in infections of bone, joints and the central
CC   nervous system, however it is particularly involved in the inflammatory
CC   lesions associated with acne vulgaris. A method for detecting the
CC   presence or absence of P. acnes in a patient comprises contacting a
CC   sample with a binding agent that binds to the proteins of the invention
CC   and determining the amount of bound protein in the sample. The
CC   polypeptides may be used as antigens in the production of antibodies
CC   specific for P. acnes proteins. These antibodies can be used to
CC   downregulate expression and activity of P. acnes polypeptides and
CC   therefore treat P. acnes infections. The antibodies may also be used as
CC   diagnostic agents for determining P. acnes presence, for example, by
CC   enzyme linked immunosorbent assay (ELISA).
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   416 AA;

  Query Match              3.3%;  Score 115.5;  DB 22;  Length 416;
  Best Local Similarity   19.8%;  Pred. No. 0.0047;
  Matches  [illegible];  Conservative   44;  Mismatches  133;  Indels  239;  Gaps   20;

Qy         47 IPPGLYHACLASLSILVLLLLAMLVRRRQLWPDCVRGRPGLPSPVDFLAGDRPRAVPAAV 106
Db          7 iqpgryqqaiqsltswal[illegible] ----eitadf---rhaltsp --- -------pwiv 47

Qy        107 FMVLLSSLCLLLPDEDALPFLTLASAPSQDGKTEAPRGAWKILGLFYYAALYYPLAACAT 166
Db         48 ssyilsdvilll---                                     -acqp 63

Qy        167 AGHAAAHLLGS[illegible]A[illegible]
Db         64 alwlrrhd[illegible]
Qy        207 LASLPLLL[illegible]
Db        124 lagisvaa[illegible]
Qy        2[illegible] KL[illegible]
Db        15[illegible]
Qy        3[illegible]  VSYLLAGRTIVLSILP[illegible]
Db        194 lvaaavad[illegible]
Qy        371 VLSCLL[illegible]
Db        234 v[illegible]
Qy        414 --RQALF[illegible]
Db        294 ilrtqiqpw[illegible]
Qy        431 PCLGLLV[illegible]
Db        35[illegible]

RESULT 11
ABB67807
ID   ABB67807 standard; protein; [illegible] AA.
XX
AC   ABB67807;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID [illegible]
XX
KW   Drosophila; developmental biology; cell s[illegible]
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; [illegible]
XX
PR   23-MAR-2000; [illegible]
PR   11-JUL-2000; [illegible]
XX
PA   (PEKE-) PE [illegible]
XX
PI   Venter JC,  A[illegible],  [illegible] PW[illegible],  Myers EW;
XX
DR   WPI; 2001-[illegible]
DR   N-PSDB; ABL[illegible]
XX
PT   New isolated [illegible] detection re[illegible]
PT   genes from [illegible]
PT   interactions
XX
PS   Disclosure; [illegible]
XX
CC   The invention [illegible]
CC   capable of detecting [illegible] or more genes [illegible]
CC   useful in deve[illegible]
CC   cell-cell interact[illegible]
CC   insecticides, [illegible] and pharmaceut[illegible]
CC   discloses [illegible] DNA sequences (AB[illegible]
```

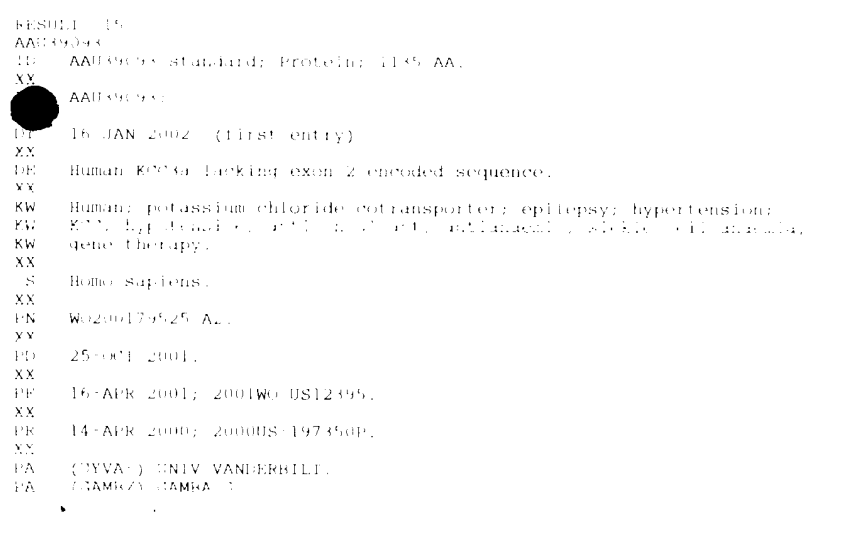
```
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   1007 AA;

  Query Match              3.8%;  Score 114.5;  DB 22;  Length 1007;
  Best Local Similarity   18.2%;  Pred. No. 0.023;
  Matches  127;  Conservative   85;  Mismatches  208;  Indels  277;  Gaps   20;

Qy         34 LQPEGEVPSCHISIPPGLYHACLASLSILVLLLLAMLVR-RRQLWPR[illegible]VRGRPGLPSPVD 92
Db        312 lqpekp[illegible] 355

Qy         93 FLAGDRPRAVPAAVFMVL---LSSLCLLLPDEDALPFLTLASAPSQDGKTEAPRGAWKI 148
Db        356 [illegible] 414

Qy        149 LGLFYYAALYYPLAACATAGHTAAHLLGSILSWAHLGVQVWQRAECP-------QVPK 199
Db        415 hwmfqm[illegible]-[illegible]---itrnvndpdtspiqemihvpi 464

Qy        200 [illegible] 248
Db        465 [illegible] 524

Qy        249 [illegible]------WARV[illegible]----------HC[illegible]YTPQP 288
Db        525 tylnnl [illegible] 581

Qy        289 GFHLP ---------------------------------------------------- 293
Db        582 epeipaadeeqvenaadenenqeqeaadaaaqppappppppppvepapaprnlaahqai 641

Qy        294 -------------------------LKLVLSATLIGTA-IYQVALLLLVGVVPTI 322
Db        642 meqrdap[illegible]ltvpvyl[illegible]rlmliwssqpd 701

Qy        323 QKVRAGV[illegible]LVVF 
```

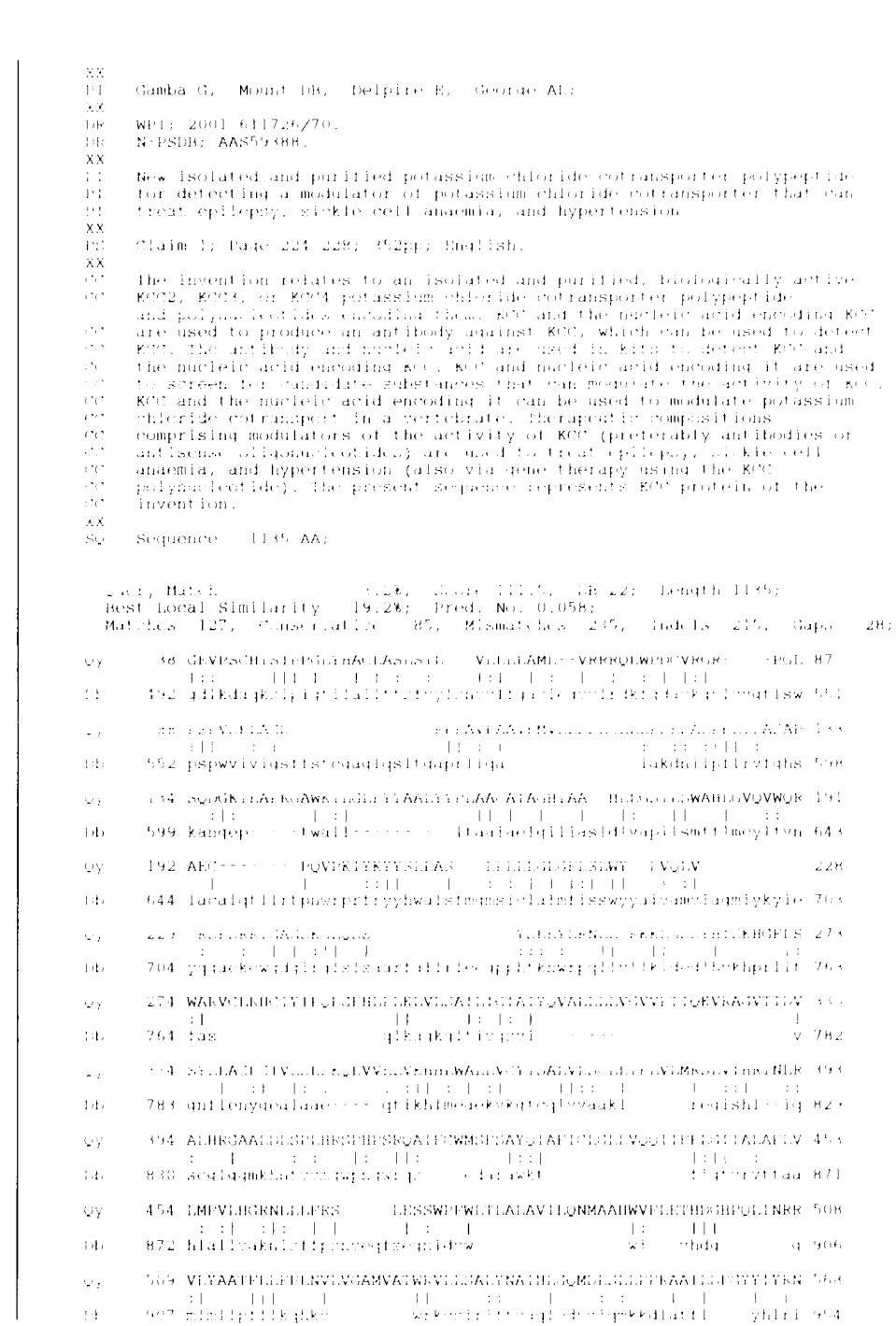
```
                                                             [illegible]WALH 363
Db        702 ekaa-----------aaaaaeqaaevvenldp[illegible]rkner[illegible]ytae[illegible] 747

Qy        364 -VCYISALVLSCLLTPLVLMRSLVTHRTN                       LRALHRGAAL 401
Db        748 ylewiv[illegible] 807

Qy        402 DLS------PLHRSP-HPSRQATFCWMSFSAYQTAFICLGLLVQQ[illegible]TFFLGT ATAFLV 453
Db        808 elavvip[illegible]rkt[illegible]ph-------tlwodwa-------lqvlyokl------a[illegible]tm 846

Qy        454 LMPVLHGRNLLLPR[illegible]W                    [illegible]LALAV[illegible]LQNMAAHWVF 494
Db        847 mqpdwh---lkr[illegible]lp 901

Qy        495 LETHDGHPQLLNRRVLYAATP              LLPPLNVL 523
Db        902 itladayaq[illegible]ipp[illegible]qsie[illegible] 949


RESULT 12
ABB62839
ID   ABB62839 standard; Protein; 423 AA.
XX
AC   ABB62839;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 15309.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL06942.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 15309; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL[illegible]), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   423 AA;

  Query Match              3.8%;  Score 111;  DB 22;  Length 423;
  Best Local Similarity   18.8%;  Pred. No. 0.0069;
  Matches   72;  Conservative   40;  Mismatches  120;  Indels  150;  Gaps   13;

Qy        241 [illegible]            [illegible]SWARVC---          278
Db         87 [illegible]wsds 136

Qy        279 ---LRHCIYTPQPGFHLPLKLV--------LSATLIGTATYQVALLLLVGVVPTIQKV 325
Db        137 ntllrkllnp[illegible]lamelmvsqaedlvri 196

Qy        326 [illegible]VVL[illegible]HHLW----------------AL 362
Db        197 [illegible]---------llaatlwlns[illegible]tqtyiwk[illegible]trvtdar 248

Qy        363 [illegible]AALD[illegible]WMS 422
Db        249 ky[illegible]ysayawqet[illegible]atme[illegible]qwla 296

Qy        423 FSAYQTAFICLGLLVQQTTFTLGTTALATLVLMPVLHGRNLLLPRGLESGWPFWLILALA 482
Db        297 [illegible]tapiact-ilvn  ittyvttr-------------------- 319

Qy        483 VILQNMAAHWVFLETHDGHPQLLNRRVLYA       ATPLLPPLNVLVIA      MVAT 530
Db        320 -----------------klinr[illegible]vyqriahklkantimtslmllvmsiawlilims 359

Qy        531 WRVLLSALYNATHLGQMDLSLL 552
Db        360 wlqmeqllyahivvnalqtpll 381
```

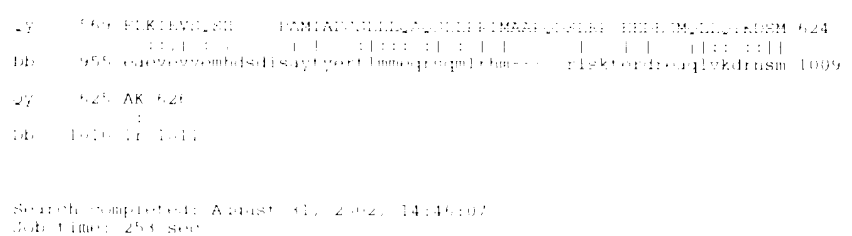
```
RESULT 13
AAU38946
ID   AAU38946 standard; Protein; 423 AA.
XX
AC   AAU38946;
XX
DT   16-JAN-2002  (first entry)
XX
DE   Drosophila G-protein coupled receptor, GPCR #24.
XX
KW   Drosophila; G protein coupled receptor; GPCR; insecticide; diagnostic;
KW   mutation detection.
XX
OS   Drosophila melanogaster.
XX
PN   WO200170980-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09341.
XX
PR   23-MAR-2000; 2000US-191638P.
PR   18-JUL-2000; 2000US-0618893.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Gravchik A;
XX
DR   WPI; 2001-[illegible]/71.
DR   N-PSDB; AAS57118, AAS57119.
XX
PT   Sixty six Drosophila Melanogaster G-protein coupled receptors (GPCR),
PT   useful in the treatment and diagnosis of GPCR-related conditions and
PT   for identifying GPCR modulators for use as insecticides -
XX
PS   Claim 1; Page 143-144; 392pp; English.
XX
CC   The invention relates to sixty six novel isolated Drosophila
CC   melanogaster G-protein coupled receptors (GPCR). The GPCR proteins and
CC   nucleic acids are useful in the treatment and diagnosis of GPCR-related
CC   conditions. The GPCR proteins and nucleic acids are also useful for
CC   identifying modulators of GPCR proteins for use as insecticides. The
CC   nucleic acid can also be used to detect mutations in GPCR genes and gene
CC   expression products such as mRNA. AAU38923-AAU38988 represent D.
CC   melanogaster G coupled protein receptor amino acid sequences of
CC   the invention.
XX
SQ   Sequence   423 AA;

  Query Match             3.3%;  Score 114;  DB 22;  Length 423;
  Best Local Similarity   18.8%;  Pred. No. 0.0069;
  Matches   72;  Conservative   43;  Mismatches  120;  Indels  150;  Gaps   13;

Qy        24? [illegible]                    EYLENLLIKKELGSSYHLGKHCFLSWAFVC    278
Db         80 [illegible]                                                   136

Qy        279    LRHCIYIPQPGFHLPLKLV            --LSATLTGIAIYQVALLLLVGVVPTIQKV 325
Db        137 ntllrkilnpithqislvillviaiiytilptlsrdlvqnivttlamclmvsqaadlvri 196

Qy        326 FA[illegible]KQEVVELVKHHLW                            AL 362
Db        197 ltelts[illegible]ts        ----llaatfwlnstqtyiwkttrsrnvtl[illegible] 248

Qy        363 EV[illegible]LMRGLVIHRINLKALHRGAALDLSPLHRSPHPSRQAIPCWMS 422
Db        249 kycyynayaw[illegible]tatmaal---------avtanftldaesyxqchmvqeqetiawlq 296

Qy        423 FSAYQIAF[illegible]IALAPLVLMPVLHGRNLLLP[illegible]WPFWLLLALA 482
```

```
Db        297 leittapl[illegible]

Qy        483 VILQNMAAHWV[illegible]INFFVLYA [illegible]

Db        326 ------ [illegible]

Qy        543 WRVLLSAYNA[illegible]

Db        36? w[illegible]qmeq[illegible]

RESULT 14
AAU39099
ID   AAU39099 standard; [illegible] AA.
XX
AC   AAU39099;
XX
DT   16-JAN-2002  [illegible]
XX
DE   Human potassi[illegible] transporter [illegible]
XX
KW   Human; potass[illegible] transporter [illegible]
KW   KCC; hypotens[illegible]
KW   gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO200179525-A.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; [illegible]
XX
PR   14-APR-2000; [illegible]
XX
PA   (UYVA-) UNIV VAN[illegible]
PA   (GAMB/) GAMBA [illegible]
XX
PI   Gamba G, Moun[illegible] A.
XX
DR   WPI; 2001-6[illegible]
DR   N-PSDB; AAS5[illegible]
XX
PT   New isolated [illegible] potassium [illegible]
PT   for detecting [illegible] potassium [illegible]
PT   treat epilepsy, [illegible] anaemia, and [illegible]
XX
PS   Claim 1; Page [illegible] English.
XX
CC   The invention [illegible] isolated and [illegible]
CC   KCC2, KCC3, or K[illegible] potassium chloride [illegible]
CC   and polynucle[illegible] them. K[illegible]
CC   are used to pr[illegible] antibody against K[illegible]
CC   KCC. The antibody and nucleic acid are used [illegible]
CC   the nucleic acid encoding K[illegible] and [illegible]
CC   to screen for candidate substances that [illegible]
CC   KCC and the nucleic acid encoding it can be [illegible]
CC   chloride cotransport in a vertebrate. [illegible]
CC   comprising modulators of the activity of K[illegible]
CC   antisense oligonucleotides) are used to [illegible]
CC   anaemia, and hypertension (also via [illegible]
CC   polynucleotide). The present sequence [illegible]
CC   invention.
XX
SQ   Sequence   [illegible] AA

  Query Match             [illegible]
  Best Local Similarity   [illegible]  Pred. N[illegible]
  Matches  127;  [illegible]  Mismatches [illegible]

Qy        [illegible]
```









Search Completed: August 31, 2002, 14:46:02
Job Time: 253 sec

OM protein - protein search, using sw model

Run on: August 31, 2002, 14:43:49 ; Search time 14.64 Seconds
(without alignments)
1112.833 Million cell updates/sec

Title: US-09-759-056-2
Perfect score: 3478
Sequence: 1 MSSQPAGNQTSPGATEDYSY..........NPTLQVFRKTALLGANGAQP 667

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: [illegible] seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: [illegible]
2: [illegible]
3: [illegible]
4: [illegible]
5: [illegible]
6: [illegible]

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

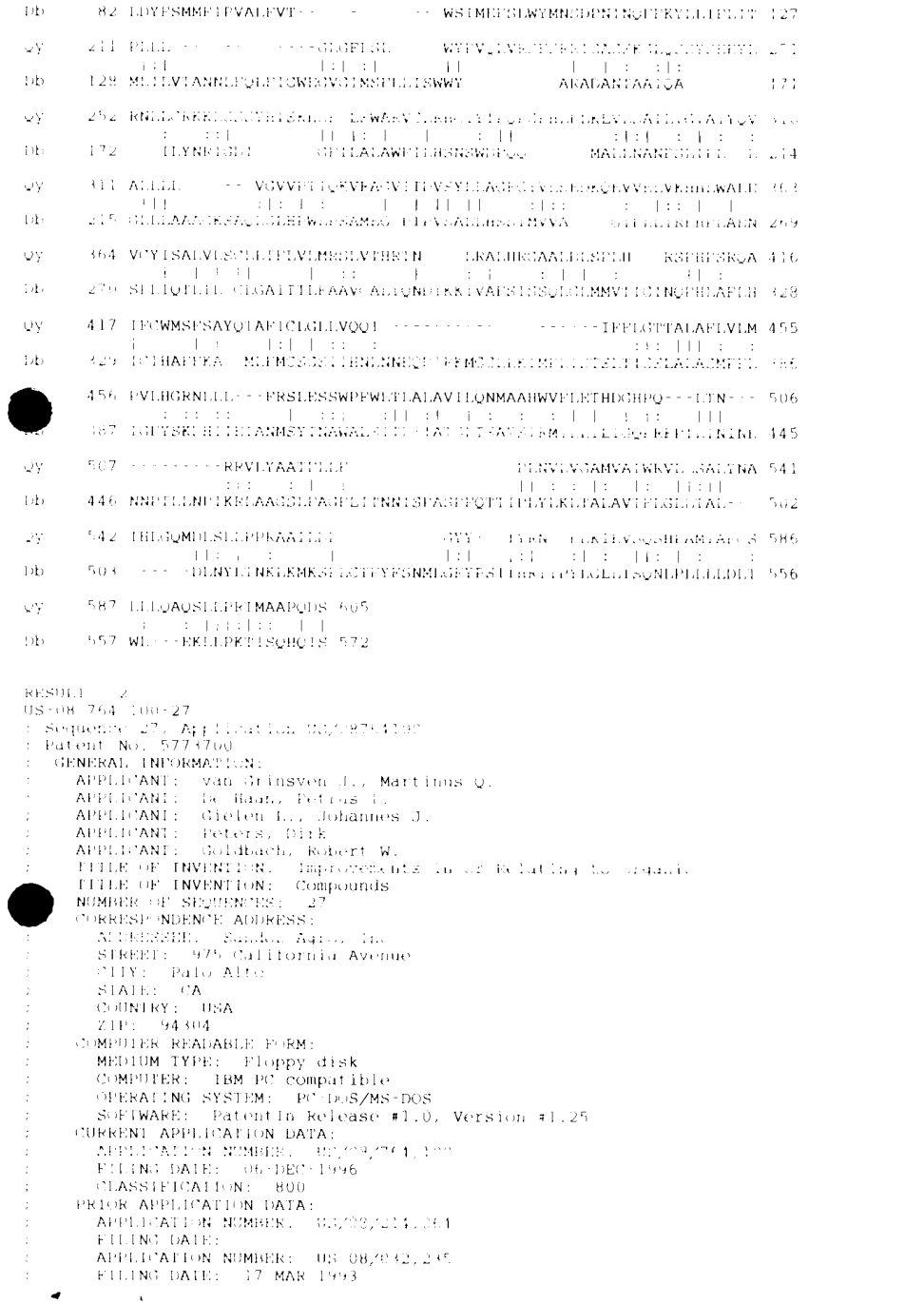
Result No.	Score	% Query Match	Length	DB	ID	Description
1	115	3.3	633	4	US-09-097-889-23	Sequence 23, Appl
2	110.5	3.2	3218	1	US-08-764-100-27	Sequence 27, Appl
3	102.5	2.9	459	4	US-09-097-889-22	Sequence 22, Appl
4	101	2.9	416	3	US-08-858-876A-4	Sequence 4, Appli
5	101	2.9	416	4	US-09-472-880-4	Sequence 4, Appli
6	100.5	2.9	2887	4	US-08-462-467B-8	Sequence 8, Appli
7	98.5	2.8	797	2	US-08-700-013B-19	Sequence 19, Appl
8	98.5	2.8	799	2	US-08-700-013B-27	Sequence 27, Appl
9	97.5	2.8	797	4	US-09-182-728A-2	Sequence 2, Appli
10	96.5	2.8	797	2	US-08-700-013B-21	Sequence 21, Appl
11	96	2.8	1495	4	US-08-462-467B-12	Sequence 12, Appl
12	95.5	2.7	312	1	US-08-118-270-38	Sequence 38, Appl
13	95.5	2.7	312	5	PCT-US93-08528-38	Sequence 38, Appl
14	95.5	2.7	355	1	US-07-759-568-1	Sequence 1, Appli
15	95.5	2.7	355	1	US-08-450-393A-8	Sequence 8, Appli
16	95.5	2.7	355	2	US-08-390-000A-5	Sequence 5, Appli
17	95.5	2.7	355	4	US-08-446-669-8	Sequence 8, Appli
18	95.5	2.7	355	5	PCT-US95-09476-8	Sequence 8, Appli
19	95.5	2.7	360	1	US-08-202-056-7	Sequence 7, Appli
20	95	2.7	315	1	US-07-866-979-4	Sequence 4, Appli
21	95	2.7	315	2	US-08-466-936B-4	Sequence 4, Appli
22	95	2.7	315	3	US-08-706-281A-4	Sequence 4, Appli
23	95	2.7	315	4	US-09-201-746-4	Sequence 4, Appli
24	95	2.7	315	4	US-09-097-231-4	Sequence 4, Appli
25	94.5	2.7	673	4	US-09-091-725-13	Sequence 13, Appl
26	94.5	2.7	673	4	US-09-091-725-19	Sequence 19, Appl
27	94.5	2.7	673	4	US-09-091-725-23	Sequence 23, Appl
28	94.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
29	94	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
30	93.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
31	93.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
32	93	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
33	92	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
34	92	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
35	92	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
36	91	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
37	91	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
38	91	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
39	91	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
40	91	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
41	91	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
42	91	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
43	91	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
44	91	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
45	91	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

ALIGNMENTS

RESULT 1
US-09-097-889-23
; Sequence 23, Application [illegible]
; Patent No. 621[illegible]
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; APPLICANT: [illegible]
; TITLE OF INVENTION: [illegible]
; TITLE OF INVENTION: [illegible]
; TITLE OF INVENTION: [illegible]
; NUMBER OF SEQUENCES: [illegible]
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: [illegible]
; STREET: [illegible]
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: [illegible]
; COMPUTER: [illegible]
; OPERATING SYSTEM: [illegible]
; SOFTWARE: [illegible]
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible]
; CLASSIFICATION: [illegible]
; ATTORNEY/AGENT INFORMATION:
; NAME: [illegible]
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: [illegible]
; TELEFAX: [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]
; SEQUENCE CHARACTERISTICS:
; LENGTH: [illegible]
; TYPE: amino acid
; STRANDEDNESS: [illegible]
; TOPOLOGY: [illegible]
US-09-097-889-23

Query Match [illegible] Score [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Mismatches [illegible]

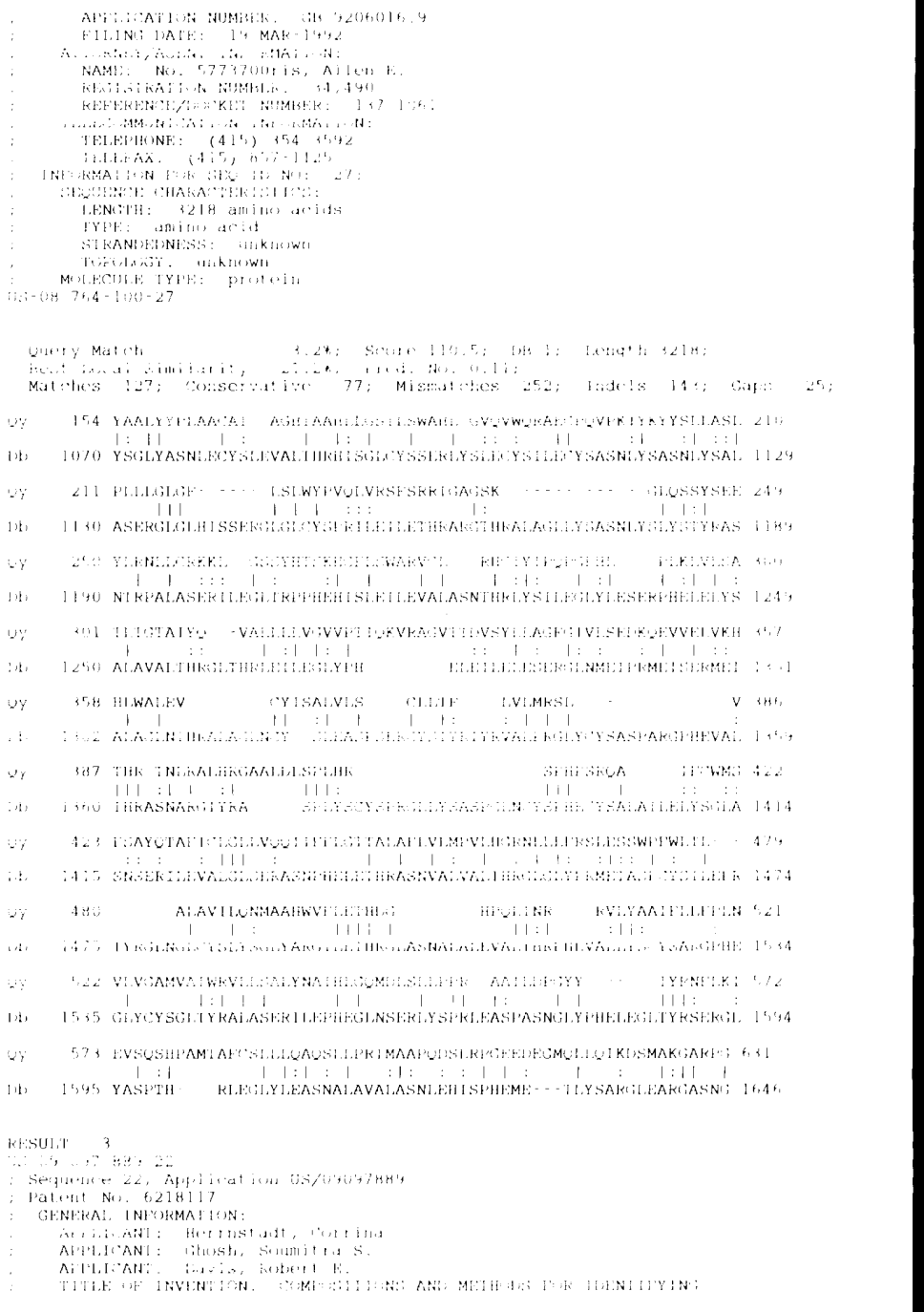
Qy [illegible] LFYYAA[illegible]AA[illegible]AAHL[illegible]SWA[illegible]



RESULT 2
US-08-764-100-27
; Sequence 27, Application US/08764100
; Patent No. 5774700
; GENERAL INFORMATION:
; APPLICANT: van Grinsven F., Martinus Q.
; APPLICANT: De Baan, Petrus T.
; APPLICANT: Gielen L., Johannes J.
; APPLICANT: Peters, Dirk
; APPLICANT: Goldbach, Robert W.
; TITLE OF INVENTION: Improvements in or Relating to Vitamin
; TITLE OF INVENTION: Compounds
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fenwick & West LLP
; STREET: 975 California Avenue
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/764,100
; FILING DATE: 06-DEC-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,251
; FILING DATE:
; APPLICATION NUMBER: US 08/032,235
; FILING DATE: 17 MAR 1993

; APPLICATION NUMBER: GB 9206016.9
; FILING DATE: 19 MAR 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 27737001S, Allen E.
; REGISTRATION NUMBER: 34,490
; REFERENCE/DOCKET NUMBER: 137 1761
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 354 4592
; TELEFAX: (415) 857-1125
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3218 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-764-100-27

Query Match 3.2%; Score 110.5; DB 1; Length 3218;
Best Local Similarity 21.2%; Pred. No. 0.11;
Matches 127; Conservative 77; Mismatches 252; Indels 144; Gaps 25;



```
;    TITLE OF INVENTION:  AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
;    TITLE OF INVENTION:  EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIO
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/097,889
;      FILING DATE:  15-JUN-1998
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rosenman Ph.D., Stephen J.
;      REGISTRATION NUMBER:  43,058
;      REFERENCE/DOCKET NUMBER:  660088.417
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  459 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-09-097-889-22

  Query Match             2.9%;  Score 102.5;  DB 4;  Length 459;
  Best Local Similarity   20.6%;  Pred. No. 1.036;
  Matches   77;  Conservative   50;  Mismatches   99;  Indels  147;  Gaps   18;

Qy    201 YKYYSLLASLPLLLGLCFLSLWYPVQLVRSFSRRICACSKCLQSSYSEEYLRNL----    254
Db    149 FLFYILVGSLPLLIAL----------------------------IYIHNILGSLNILLLL 180

Qy    255 LCRKKLCSSTHISKHFGFLSWARVCLRHTYTFQPGFHLPL KLVLSATLLIGIAIYQVALL 313
Db    181 LIAQELGNGWANN     LMWLAYIMAFMVKMFLYGLHLWLFKAFVEAFIAGSMVLAAVLL 236

Qy    314 LLVGVVFILQKVRAGVILCVSTLLAGFGIVLSEDKQEVVELVKEHLWALEVCYISALVLS 373
Db    237 KLGG--            ---YGMMRLILILN     ---PLIKHMAYPF-------LVLS 265

Qy    374 CLLTFLVLMRCCVTHRPNLFALHRCAALFLSFLLEGTHLGRLAHLWMCTGAYQIAFICL 433
Db    266  LWGMIMISSICLRQIDLKSL ----------         --- ----IAYSSIS--HMALVVT 299

Qy    434 GLLVGQTTFFLG---- ----TTALAFLVL    MPVLHGRNLLLFRGLESSWH    FWL 477
Db    300 AIFLGIPWTPTGAVIFMIASYIPCILPFLANGNYEEIHGFIM LSQCLQILLPLMAFWW 359

Qy    478 TLALAVILLNMAAHWVFLETHQGHPQLLNRRVLYAATFLLFPLNVIVGAMVAIW----   531
Db    360 LLA    SLANLAL            -PPLIN-          LLGELSVLVTTF--SWSNITLL 394

Qy    532      RVLLSALYN 540
Db    395 LLGLNMLVIALYS 407

RESULT  4
US-08-858-876A-4
; Sequence 4, Application US/08858876A
; Patent No. 6022856
;  GENERAL INFORMATION:
```

```
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    TITLE OF INVENTION: [illegible]
;    TITLE OF INVENTION: [illegible]
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: [illegible]
;      STREET: [illegible]
;      CITY:  Washington
;      STATE: [illegible]
;      COUNTRY: [illegible]
;      ZIP: [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: [illegible]
;      COMPUTER: [illegible]
;      OPERATING SYSTEM: [illegible]
;      SOFTWARE: [illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;      CLASSIFICATION: [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Player, William [illegible]
;      REGISTRATION NUMBER: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible]
;      TYPE: [illegible]
;      TOPOLOGY: [illegible]
;    MOLECULE TYPE: [illegible]
US-08-858-876A-4

  Query Match  [illegible]  Score [illegible]
  Best Local Similarity  [illegible]  Pred. No. [illegible]
  Matches   86;  Conservative [illegible]  Mismatches [illegible]

Qy    285 TPQPGFH[illegible]KLVLSATLL[illegible]IAIYQVALL[illegible]
Db     12 SPSAGLS[illegible]WAFVIFIALY[illegible]
Qy    335 YLLAGF[illegible]VKEHLWALEV[illegible]
Db     [illegible] Y[illegible]
Qy    395 LHRGAAL[illegible]WMS[illegible]SAY[illegible]
Db     88 [illegible]YN[illegible]VW[illegible]
Qy    451 FL[illegible]
Db    128 SLSAER[illegible]
Qy    500 GHPQLIN[illegible]NV[illegible]
Db    185 GEPEFASEV[illegible]VNVLVS[illegible]
Qy    547 M[illegible]YY[illegible]YRN[illegible]
Db    245 AQVSS[illegible]WRK[illegible]

RESULT  5
US-09-472-880-4
; Sequence 4, Application [illegible]
; Patent No. 627[illegible]
;  GENERAL INFORMATION:
;       APPLICANT: [illegible]
```

```
;             Pascale CHALON
;             Pascual FERRARA
;             Vita NATALIO
;       TITLE OF INVENTION: Type 2 Neurotensin Receptor
;                           (hNT-R2)
;       NUMBER OF SEQUENCES: 12
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
;            STREET: 400 Seventh Street
;            CITY: Washington
;            STATE: D.C.
;            COUNTRY: USA
;            ZIP: 20004
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/472,880
;            FILING DATE: 28 Dec 1999
;            CLASSIFICATION: Unknown
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: PCT/FR 97/23283
;            FILING DATE: 17 MAR 1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Player, William E.
;            REGISTRATION NUMBER: 31,049
;   INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 416 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-472-880-4

  Query Match             2.9%;  Score 101;  DB 4;  Length 416;
  Best Local Similarity   24.2%;  Pred. No. 0.049;
  Matches   86;  Conservative   45;  Mismatches  122;  Indels  112;  Gaps   18;

Qy        [illegible]
Db         12 SPSAGLSLEARLGVDTRLWAKVLFTALY--SLIFAFGIAGNALSVHVVLKARAGRPGRLR 69

Qy        335 YLLAGF[illegible] 394
Db         70 Y------------------------------HVLSLALSALLLLLVSM------------ 87

Qy        [illegible]
Db         88 --------------PMELYNFVW  [illegible] 127

Qy        451 FL----VLMFVLHGRNL[illegible] 499
Db        128 SLSA[illegible] 184

Qy        500 GHPQLINR  RVLYA ATFLLF FLNVLVG      AMVAIWRVL     LSALYNA[illegible] 546
Db        185 GEPEPASRV[illegible] 244

Qy        547 MDLSLLPPRAATLDP     GYYTYRNFLKIEV[illegible]   HPAMTAP[illegible] 594
Db        245 AQV[illegible]

RESULT 6
US-08-462-467B-8
; Sequence 8, Application US/08462467B
; Patent No. 6210899
;  GENERAL INFORMATION:
;    APPLICANT: Rosenbaum, Jan S
;    TITLE OF INVENTION: The Use of a BMP Protein Receptor
;    TITLE OF INVENTION: Complex for Screening Bone Metabolism Actives and Cells
;    TITLE OF INVENTION: Co-Transfected With a Type II BMP Receptor and a Type I
;    TITLE OF INVENTION: BMP Receptor
;    NUMBER OF SEQUENCES: 39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: The Procter & Gamble Company
;      STREET: 11810 East Miami River Road
;      CITY: Ross
;      STATE: OH
;      COUNTRY: USA
;      ZIP: 45061
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/462,467B
;      FILING DATE:
;      CLASSIFICATION: 435
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Hersko, Bart S.
;      REGISTRATION NUMBER: 32,572
;      REFERENCE/DOCKET NUMBER: 5174R
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (513) 627-0633
;      TELEFAX: (513) 627-0260
;  INFORMATION FOR SEQ ID NO: 8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 2887 amino acids
;      TYPE: amino acid
;      STRANDEDNESS: not relevant
;      TOPOLOGY: not relevant
;    MOLECULE TYPE: protein
US-08-462-467B-8

  Query Match             2.9%;  Score 100.5;  DB 4;  Length 2887;
  Best Local Similarity   21.9%;  Pred. No. 1.1;
  Matches  123;  Conservative   53;  Mismatches  187;  Indels  199;  Gaps   28;

Qy        [illegible] 197
Db        856 THRARG     GLYLEALATYFLEHHSTHPGLLFFRAPGGYASPHIST-      ---  - 901

Qy        [illegible] 257
Db        902  -YRLYS-----PRALAT---  ---LESERHISAR--GASPLEASNSE         933

Qy        [illegible] VALLLLVG 317
Db        934 KARGASN        VALLEVALLYSASNASEG             LYALACYSVALIL     971

Qy        [illegible] 377
Db        972      [illegible]  [illegible]GLYLESGRMETA         993

Qy        378 [illegible]                 [illegible]WMSFSAYQI 428
Db        [illegible]  [illegible]AGNALAA LATLESER 1044

Qy        429 AF[illegible]  [illegible]WLTLALAVI 484
Db        [illegible] 1089

Qy        485 LQNMAAHWVFLETHDGHPQLINRRVLYAATFLLFFLNV LVGAMVAIWRVL LSALYNAI 542
Db       1090 LASN-     --LEARGASPGYSGLSERALALELYSGLNVALASPMETTYRALALEGLYLET 1143

Qy        543 ---HLSQMDLSLLP-       --    --PRAATLDP  --GYY        --      TYRNFL 575
Db       1144 LETYRTRPGLVALPHEMETARGGYSTHRASPLEPHEPRGLYGLSERVALPRASPTYRG L 1202
```

```
Qy    5?? KIEVS?? HPAMIA???      ?QA?S??PFIMAAP??SLR???????M?????KRCMA 625
Db   12?? NMSTALALH?????HF??VAL??YASN???KTHKF???LAS?ML???NVALLEV-       1257

Qy    626 KGARPGASRGEARWGLAYIL?R 647
Db   1258        ALSEFARGGLLYS?LN 127?
```

RESULT 7
US-08-700-013B-19
; Sequence 19, Application US/08700013B
; Patent No. 5919653
; GENERAL INFORMATION:
;   APPLICANT: Albert, Vivian R.
;   APPLICANT: Kowalski, Leslie R.Z.
;   APPLICANT: Borden, Laurence A.
;   APPLICANT: McKelvy, Jeffrey F.
;   TITLE OF INVENTION: Human Glycine Transporter
;   NUMBER OF SEQUENCES: 41
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Dechert Price & Rhoads
;   STREET: 997 Lenox Drive, Building 3, Suite 210
;   CITY: Lawrenceville
;   STATE: NJ
;   COUNTRY: USA
;   ZIP: 08543
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/700,013B
;   FILING DATE:
;   CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Bloom, Allen
;   REGISTRATION NUMBER: 29,135
;   REFERENCE/DOCKET NUMBER: 317743-108
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 609-520-3214
;   TELEFAX: 609-520-3259
;   TELEX:
; INFORMATION FOR SEQ ID NO: 19:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 797 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-08-700-013B-19

  Query Match   2.8%; Score 98.5; DB 2; Length 797;
  Best Local Similarity   20.9%; Pred. No. 0.23;
  Matches   88; Conservative   57; Mismatches 168; Indels 109; Gaps   23;

```
Qy    139 IEAPR?AWK              ??GLFYYAALYYFLAACATAGHTAAHLLGSILSWAHLGVQ 187
Db    25? SQ?PVSVWR????AL?????AML???VLIAIYYNVI?CYILFYLFA?FV SVLPWG??N N 31?

Qy    188 VWQRAE??QVPKIYKYYSLLASLP  LL??GLGFLSLWYPVQLVRSFSRRIGAGSKGLQSS 245
Db    314 FWN??? ?K?F?F?????V???????????NSTF??TATFNVIMVNF    ??QANKIFV?? 370

Qy    246 Y?FLYL?NLL?RKR???S???????SWARV?LRP??Y------- IPQPCPHLPLKLV 297
Db    37?  SEAYLKYIVLKIGA???      ???IRWI    LAL?LPLAWVIVYASLAKGIKLGKVV 422

Qy    298 LSAILI?IA?????????VVFV??KVEA??
Db    42? Y???FIA???????????IE???????????
Qy    337 LAGFG  ???????????????WALFV???????
Db    47? SAAWG?????      ???   NN????
Qy    388 HFINLRA????AA?????          ????
Db    52? FMANERKVN??NVA???????????????????
Qy    43? LVQQI          ???????????AEIVLM???
Db    58? IVISIS?????????????????????????
Qy    48? VI 484
Db    64? II 644
```

RESULT 8
US-08-700-013B-27
; Sequence 27, Application US/08700013B
; Patent No. 59196??
; GENERAL INFORMATION:
;   APPLICANT: Albert, Vivian R.
;   APPLICANT: Kowalski, Leslie R.Z.
;   APPLICANT: Borden, Laurence A.
;   APPLICANT: McKelvy, Jeffrey F.
;   TITLE OF INVENTION: Human Glycine Transporter
;   NUMBER OF SEQUENCES: 41
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Dechert Price & Rhoads
;   STREET: 997 Lenox Drive, Building 3, Suite 210
;   CITY: Lawrenceville
;   STATE: NJ
;   COUNTRY: USA
;   ZIP: 08543
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/700,013B
;   FILING DATE:
;   CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Bloom, Allen
;   REGISTRATION NUMBER: 29,135
;   REFERENCE/DOCKET NUMBER: 317743-108
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: [illegible]
;   TELEFAX: [illegible]
;   TELEX:
; INFORMATION FOR SEQ ID NO: 27:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 797 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
US-08-700-013B-27

  Query Match   [illegible]
  Best Local Similarity   [illegible]
  Matches   88; [illegible] Mismatches [illegible]

```
Qy    139 IEAPRGAWK              ??LFYYAALYY?AA?? [illegible]
```

RESULT 9
US-09-182-728A-2
; Sequence 2, Application US/09182728A
; Patent No. 6248883
; GENERAL INFORMATION:
; APPLICANT: BROWN, ANTHONY
; APPLICANT: CHAPMAN, CONRAD GERALD
; APPLICANT: GLASER, ISRAEL SIMON
; APPLICANT: EVANS, JOANNE RACHEL
; APPLICANT: CAIRNS, WILLIAM
; APPLICANT: HERDON, HUGH
; TITLE OF INVENTION: NOVEL P MEMBRANE
; FILE REFERENCE: GP 30176
; CURRENT APPLICATION NUMBER: US/09/182,728A
; CURRENT FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 9818890.7
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 797
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-182-728A-2

Query Match 2.8%; Score 97.5; DB 4; Length 797;
Best Local Similarity 20.6%; Pred. No. 0.29;
Matches 87; Conservative 58; Mismatches 168; Indels 109; Gaps 23;

RESULT 10
US-08-700-013B-21
; Sequence 21, Application US/08700013B
; Patent No. 5919653
; GENERAL INFORMATION:
; APPLICANT: Albert, Vivian R.
; APPLICANT: Kowalski, Leslie R.Z
; APPLICANT: Borden, Laurence A.
; APPLICANT: McKelvy, Jeffrey F.
; TITLE OF INVENTION: Human Glycine Transporter
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beehert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 08/700,013B
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen
; REGISTRATION NUMBER: 29,135
; REFERENCE/DOCKET NUMBER: 21741-108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609 520 3214
; TELEFAX: 609-520-3259
; TELEX:
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 797 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-700-013B-21

Query Match 2.8%; Score 96.5; DB 2; Length 797;
Best Local Similarity 20.6%; Pred. No. 0.29;
Matches 87; Conservative 56; Mismatches 170; Indels 109; Gaps 23;

```
Qy   139 IEAPRGAWKII[illegible]YYAALYYPLAA[illegible]AIACHIAAHI[illegible]WAHLGVQ 187
Db   256 [illegible]
Qy   186 VWQRAI[illegible]QVIKIYKYYSLLASLI[illegible]LLG[illegible]GFLSLWYPVQLVRSFSRRTGAGSKGLQSS 245
Db   314 PWNIPE[illegible]IFNTIF[illegible]MVN[illegible]ANRIIV[illegible] 370
Qy   246 YSEEYLRNL[illegible]RKKL[illegible]SYHISKH[illegible]LSWARVCLRH[illegible]IY--------IPQPGFHLPLKLV 297
Db   371 SEEYIF[illegible]VIF[illegible]FW---LAI[illegible]LAWV[illegible]VYASLARG[illegible]KVV 422
Qy   298 LSA[illegible]NIY[illegible]VALLL[illegible]VVF[illegible]KVEA[illegible]V[illegible] 336
Db   423 Y FIAIFFYVVLVILLI[illegible]GV ILFGAGAGIWYFI[illegible]RWEELIIA[illegible]W[illegible]DAAFQIFPSL 478
Qy   337 IA[illegible]VI[illegible]F[illegible]EVVELVEHHLWALEV[illegible]Y[illegible]V[illegible]MRSIVI 387
Db   479 SAAW[illegible]LLSSYN KFH [illegible]NNCYRDI[illegible]IVICTNSAT[illegible]IAGFVIF SVIG 525
Qy   388 HRIN[illegible]RAIHR[illegible]AAI[illegible]F[illegible]I[illegible]HEF[illegible]LAI[illegible]WMSF[illegible]AYLTAFIGL[illegible] 435
Db   526 FHANEKKVNIENVAL[illegible]IAPVVYFEAL[illegible]KLFL[illegible]FWAI[illegible]FFLML[illegible]IMFATIEI 585
Qy   436 IVQQI-------------IPPLGTIALAPLVLMPVLHGRNLLLPRGL[illegible]SWPIWLTLALA 482
Db   590 IV[illegible]PKYL[illegible]HKIVFILS[illegible]VCIFIMGFPM[illegible]YMPQLV[illegible] YAASTALV 642
Qy   483 VI 484
Db   643 II 644
```

RESULT 11
US-08-462-467B-12
; Sequence 12, Application US/08462467B
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: Rosenbaum, Jan S
; TITLE OF INVENTION: The Use of a BMP Protein Receptor
; TITLE OF INVENTION: Complex for Screening Bone Metabolism Actives and Cells
; TITLE OF INVENTION: Co-Transfected With a Type II BMP Receptor and a Type I
; TITLE OF INVENTION: BMP Receptor
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Procter & Gamble Company
; STREET: 11810 East Miami River Road
; CITY: Ross
; STATE: OH
; COUNTRY: USA
; ZIP: 45[illegible]
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,467B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hersko, Bart S.
; REGISTRATION NUMBER: 32,572
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (513) 627-0633
; TELEFAX: (513) 627-0260
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1495 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: [illegible]
US-08-462-467B-12

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible] Mismatches [illegible]

```
Qy   [illegible] [illegible]AVIAAVIMV[illegible]
Db   797 EIARGHISCLA[illegible]
Qy   [illegible] KIEAPRGAWK[illegible]YYAALY[illegible]
Db   839 [illegible]
Qy   [illegible] WAHLGVQVW[illegible]KIYKYYSLLA[illegible]
Db   887 [illegible]YSALAIHPLEA[illegible]
Qy   [illegible] GAGSKGLQ[illegible]
Db   947 SERALAALA[illegible]
Qy   282 [illegible]
Db   [illegible] LYSFRALA[illegible]
Qy   329 VIIDVSYLLA[illegible]VVELVKHHLWAL[illegible]
Db   1053 LYSERCYS[illegible]
Qy   386 VIHRINLRAIHR[illegible]AA[illegible]
Db   1100 PIHRAS [illegible]
Qy   446 I[illegible]IALA[illegible]
Db   1146 IYRMEIALAFP[illegible]VAL[illegible]ASP[illegible]SERLEASNLYSA[illegible]NHI
Qy   477 LILALAVII[illegible]NMAAB WVFLPIHH[illegible]
Db   1202 IALAASP[illegible]
Qy   513 AIFLLPF[illegible]NV[illegible]AMVAIWR VI[illegible]SALYNAIHL[illegible]
Db   1262 [illegible]LYIL[illegible]VA[illegible]
Qy   570 LK-IEVSQSE[illegible]AMIA[illegible]
Db   1317 ASPMEIAR[illegible]
Qy   621 KRGMAKGAE[illegible]
Db   1361 KASNARGIE[illegible]
```

RESULT 12
US-08-118-270-38
; Sequence 38, Application US/08118270
; Patent No. 5508[illegible]
; GENERAL INFORMATION:
; APPLICANT: Murphy, [illegible]
; APPLICANT: [illegible]
; TITLE OF INVENTION: [illegible]
; TITLE OF INVENTION: [illegible]
; NUMBER OF SEQUENCES: [illegible]
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: [illegible] & [illegible]MARK
; STREET: 419 [illegible] N.W.[illegible]
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA

```
;       ZIP:  20004
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/118,270
;       FILING DATE:  09-SEP-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/943,236
;       FILING DATE:  10-SEP-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Townsend, Kevin G.
;       REGISTRATION NUMBER:  34,033
;       REFERENCE/DOCKET NUMBER:  MURPHY 2A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-628-5197
;       TELEFAX:  202-737-3528
;       TELEX:  248633
;   INFORMATION FOR SEQ ID NO:  38:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  312 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-118-270-38

  Query Match              2.7%;  Score 95.5;  DB 1;  Length 312;
  Best Local Similarity   21.6%;  Pred. No. 0.11;
  Matches   74;  Conservative   49;  Mismatches  118;  Indels  111;  Gaps   14;

Qy        364 VCYIGALVLSIL     LTFLVLMKGLVTHRTNLKALHRGAALDLGFLAKGTHLGRQAIFCW 420
Db          6 IIYALVFLLSLLGNSLVMLVILYSRVGRSVTDVYLLNLALADLLFALTLPIWAASKVNGW 65

Qy        421 MSFSAYQTAFPTCGLLVQQIIFPIFTTALA       PLVLM     PVLHGRNLLLFRSLE 470
Db         66 IF     FGTFLCKWVSLLKEVNFY   LLLLACISVDRYLAIVHATRTLTQKRHLVKF    119

Qy        471 SSWFFWLTLALAVILQNMAAHWVFLETHDGHPQLTNRRVLYAATFLLFPLNVL        V 524
Db        120 SIWGLSLLLALPVLL---------------------FRRTVYSS          NVSPACYEDM 152

Qy        525 GAMVATWRVLLGALYNATHLLQMDLSLLPFRAATLLFGYYTTRNFLRLTVSQGRLAMIA  583
Db        153 GNNYANWRMLLRIL       FQSFGFIVPLLIMLYCYGFTLRTLFKAHMGQKHRAMRVI 205

Qy        584 ---------FC----SLLLQAQSLLPRTMAAPQDSLRPGEEDEGMQL--------------- 617
Db        206 FAVVLIFLLCWLPYNLVLLADTLMRTQVIQETCERRNHIDRALDATEILGILHSCLNPL 264

Qy        618 -------------------LQTRESMAKGARP     GASRG 635
Db        265 IYAFIGQKFRHGLLKILAIHGLISKDSLPKDSRPSFVGSSS 306

RESULT 13
PCT-US93-08528-38
; Sequence 38, Application PC/TUS9308528
;   GENERAL INFORMATION:
;     APPLICANT:  New York University
;     TITLE OF INVENTION:  POLYPEPTIDES OF G-COUPLED PROTEIN
;     TITLE OF INVENTION:  RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
;     NUMBER OF SEQUENCES:  348
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  BROWDY AND NEIMARK
;       STREET:  419 Seventh Street, N.W., Suite 300
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  USA
```

```
;       ZIP:  20004
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  [illegible]/08528
;       FILING DATE:  09-SEP-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/943,236
;       FILING DATE:  10-SEP-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Townsend, Kevin G.
;       REGISTRATION NUMBER:  34,033
;       REFERENCE/DOCKET NUMBER:  MURPHY 2 PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-628-5197
;       TELEFAX:  202-737-3528
;       TELEX:  248633
;   INFORMATION FOR SEQ ID NO:  38:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  312 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
PCT-US93-08528-38

  Query Match              2.7%;  Score 95.5;  DB 5;  Length 312;
  Best Local Similarity   21.6%;  Pred. No. 0.11;
  Matches   74;  Conservative   49;  Mismatches  118;  Indels  111;  Gaps   14;

Qy        364 VCYIGALVLSIL     LTFLVLMKGLVTHRTNLKALHRGAALDLGFLAKGTHLGRQAIFCW 420
Db          6 IIYALVFLLSLLGNSLVMLVILYSRVGRSVTDVYLLNLALADLLFALTLPIWAASKVNGW 65

Qy        421 MSFSAYQTAFPTCGLLVQQIIFPIFTTALA       PLVLM     PVLHGRNLLLFRSLE 470
Db         66 IF     FGTFLCKWVSLLKEVNFY   LLLLACISVDRYLAIVHATRTLTQKRHLVKFICL 119

Qy        471 SSWFFWLTLALAVILQNMAAHWVFLETHDGHPQLTNRRVLYAATFLLFPLNVL        V 524
Db        120 SIWGLSLLLALPVLL                     FRRTVYSS          NVSPACYEDM 152

Qy        525 GAMVATWRVLLGALYNATHLLQMDLSLLPFRAATLLFGYYTTRNFLRLTVSQGRLAMIA  583
Db        153 GNNYANWRMLLRIL       FQSFGFIVPLLIMLYCYGFTLRTLFKAHMGQKHRAMRVI 205

Qy        584 ---------FC----SLLLQAQSLLPRTMAAPQDSLRPGEEDEGMQL--------------- 617
Db        206 FAVVLIFLLCWLPYNLVLLADTLMRTQVIQETCERRNHIDRALDATEILGILHSCLNPL 264

Qy        618                    LQTRESMAKGARP     GASRG 635
Db        265 IYAFIGQKFRHGLLKILAIHGLISKDSLPKDSRPSFVGSSS 306

RESULT 14
US-07-759-568-1
; Sequence 1, Application US/07759568
; Patent No. 5374506
;   GENERAL INFORMATION:
;     APPLICANT:  Murphy, Philip M.
;     TITLE OF INVENTION:  Cloning of cDNA Encoding a Functional
;     TITLE OF INVENTION:  Human Interleukin-8 Receptor
;     NUMBER OF SEQUENCES:  5
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Cushman, Darby & Cushman
;       STREET:  1615 L Street, N.W.
;       CITY:  Washington
;       STATE:  D.C.
```

```
;       COUNTRY:  USA
;       ZIP:  20036-5601
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  
;       FILING DATE:  19910913
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Scott, Watson T.
;       REGISTRATION NUMBER:  26581
;       REFERENCE/DOCKET NUMBER:  
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-861-3000
;       TELEFAX:  202-822-0944
;       TELEX:  6714627 cush
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  355 amino acids
;       TYPE:  AMINO ACID
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-07-759-568-1

  Query Match             2.7%;  Score 95.5;  DB 1;  Length 355;
  Best Local Similarity   21.8%;  Pred. No. 0.13;
  Matches   78;  Conservative   43;  Mismatches  121;  Indels  115;  Gaps   16;

Qy        EVVELVKHHLWALEV YISALVL  L     L IFLVIMRSLVTHRIN KALHKGAALDL  L 406
Db     47 ESLEINKYF    VVIIYALVFLLSLLGNSLVMLVILYSKVGKSVIDVYLLNLALADLLFA 93

Qy    407 HRSPHPSRQAIFCWMSFSAYQIAFIC-LGLLVQQIIFFLGITALAFLVL---MPVLHG - 460
Db     94 LTLPIWAASKVNGWIF   --GTFLCKVVSLLKEVNFYSGILLLACISVDRYLAIVHATR 148

Qy    461       -RN                WPFW    AVI  NMA  WVF  IHE  HF  LINEEVLYAAIF 515
Db    149 ILTQKRYLVKFI  L-SIWGLSLLLALPVLL----                    FRRTVYSS   185

Qy    516 LLEPLNVL        VGAMVAIWRVLL ALYNAIHL QMD  SLL   AAT    GYYIYRNF 569
Db    186 -      NV  FA  YE M NNTANWRMLLRIL          F  P  IVPLLIMLF  YG  TLRTL 233

Qy    570 LKIEVS SHPAMIA-      --     F    ---SLLLQASSLLPRIMAAFQ  LRPGEEDEGMQ 616
Db    234 FKAHMG KHRAM VIFAVVLIFLL WLPYNLVLLADTLM  IQVI  ET  RRNHIDRALD 292

Qy    617 L       -  -   -   -   -   -  --  LQTKDSMAKCARP   GASRG 635
Db    293 AT                NE  YAF  KFRH  IK  A      LISKPSLPEISR  IVGSSSC 349

RESULT 15
US-08-450-393A-8
; Sequence 8, Application US/08450393A
; Patent No. 5707815
;  GENERAL INFORMATION:
;    APPLICANT:  Charo, Israel
;    APPLICANT:  Coughlin, Shaun
;    TITLE OF INVENTION:  MAMMALIAN MONOCYTE CHEMOATTRACTANT
;    TITLE OF INVENTION:  PROTEIN RECEPTORS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooley Godward Castro Huddleson & Tatum
;      STREET:  5 Palo Alto Square
;      CITY:  Palo Alto
;      STATE:  California
```

```
;       COUNTRY:  USA
;       ZIP:  94
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  
;       COMPUTER:  
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  
;       FILING DATE:  
;       CLASSIFICATION:  
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  
;       REGISTRATION NUMBER:  
;       REFERENCE/DOCKET NUMBER:  
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  
;       TELEFAX:  
;       TELEX:  
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  
;       TYPE:  amino acid
;       TOPOLOGY:  
;     MOLECULE TYPE:  
;     HYPOTHETICAL:  N
US-08-450-393A-8

  Query Match             
  Best Local Similarity   
  Matches   78;  Conservative   
  
Qy    356 EVVELVKHH
Db     47 ESLEINKYF     VVI
Qy    407 HRSPHPSRQA
Db     94 LTLPIWAASKVN         KVVSLLKEVNFY
Qy    461       RN                  AVI  NMA  WVF
Db    149 ILTQKRYLVKF            LLALPVLL
Qy    516 LLEPLNVL           ALYNAIH  QM
Db    186       NVSFA           ANWRMLLRI
Qy    570 LKIEVSQSHPAMIA          F      SLL
Db    234 FKAHMGQKH            WLPYNLVLLA
Qy    617 L
Db    293 ATE

Search completed: Aug 31 14:
Job time: 161 sec
```

GM protein - protein search, using sw model

Run on: August 31, 2002, 14:44:29 ; Search time 21.37 Seconds
(without alignments)
2999.138 Million cell updates/sec

Title: US-09-759-056-2
Perfect score: 3478
Sequence: 1 MSSQPAQNQTSPGAIELYSY.....NPTLQVFKKIALLGANGAQP 667

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3478	100.0	459	2	I11489	NADH dehydrogenase
2	2351	67.6	459	2	S41844	NADH dehydrogenase
3	2346	67.5	603	2	I44149	NADH dehydrogenase
4	1277.5	36.7	459	1	QXMS4M	NADH dehydrogenase
5	1273.5	36.6	461	2	B50885	NADH dehydrogenase
6	1270	36.5	459	2	S26169	NADH dehydrogenase
7	1259	36.2	460	2	I11309	NADH dehydrogenase
8	1225.5	35.2	459	2	S04756	NADH dehydrogenase
9	1221.5	35.1	614	2	JQ0619	NADH dehydrogenase
10	1201.5	34.5	458	2	B90623	NADH dehydrogenase
11	1200.5	34.5	459	2	I11256	NADH dehydrogenase
12	1200	34.5	459	1	QXBJ4M	NADH dehydrogenase
13	1200	34.5	460	2	I11543	NADH dehydrogenase
14	1199	34.4	461	2	I11829	NADH dehydrogenase
15	1197.5	34.4	458	2	I13819	NADH dehydrogenase
16	1197.5	34.4	605	2	S21297	NADH dehydrogenase
17	1197	34.4	460	2	I09956	NADH dehydrogenase
18	1197	34.4	602	2	I10996	NADH dehydrogenase
19	1197	34.4	518	2	JQ0198	proline/betaine tr
20	1196	34.4	518	2	A83088	MFS permease [prol
21	1195.5	34.4	459	2	I10981	NADH dehydrogenase
22	1195.5	34.4	459	2	I11202	NADH dehydrogenase
23	1195.5	34.4	459	2	A47675	NADH dehydrogenase
24	1195.5	34.4	692	2	F90680	probable ferric tr
25	1195.5	34.4	692	2	B85581	probable ferric tr
26	1195	34.4	459	2	I11333	NADH dehydrogenase
27	1195	34.4	460	2	I20866	NADH dehydrogenase
28	1194.5	34.4	459	2	I11450	NADH dehydrogenase
29	1194.5	34.4	461	1	QXXL4M	NADH dehydrogenase
30	1194.5	34.4	461	2	S70924	NADH dehydrogenase
31	1194	34.3	459	1	QXJ34M	NADH dehydrogenase
32	1194	34.3	459	2	S71844	NADH dehydrogenase
33	1193	34.3	459	2	I11340	NADH dehydrogenase
34	1193	34.3	459	2	S51194	NADH dehydrogenase
35	1193	34.3	459	2	I11774	NADH dehydrogenase
36	1192.5	34.3	459	2	S74943	NADH dehydrogenase
37	1192.5	34.3	459	2	I11130	NADH dehydrogenase
38	1192.5	34.3	459	2	A54295	NADH dehydrogenase
39	1192.5	34.3	460	2	I11554	NADH dehydrogenase
40	1192	34.3	459	2	S93564	NADH dehydrogenase
41	1192	34.3	459	2	I11394	NADH dehydrogenase
42	1192	34.3	459	2	I11 [illegible]	NADH dehydrogenase
43	1192	34.3	460	2	I14435	NADH dehydrogenase
44	1191.5	34.3	459	2	I11502	NADH dehydrogenase
45	1191.5	34.3	460	2	I12205	NADH dehydrogenase

ALIGNMENTS

RESULT 1
I11489
NADH dehydrogenase [illegible]
C;Species: mitochondrion [illegible]
C;Date: 16-Jul-1993 #sequence_revision 16-Jul-1993 #text_change [illegible]
C;Accession: I11489
R;Cissi, C.; Submitted (-- Apr-1993) to [illegible]
Genomics 50, 161 (1998)
A;Title: The complete mitochondrial DNA sequence [illegible]
A;Reference number: [illegible]
A;Accession: I11489
A;Status: preliminary; translated from GB/EMBL/
A;Molecule type: DNA
A;Residues: 1-459 <1>
A;Cross-references: EMBL:A [illegible]; NID:g1 [illegible]
C;Genetics:
A;Gene: NADH4
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: NADH dehydrogenase (ubiquinone) [illegible]
C;Keywords: membrane [illegible]; oxidoreductase [illegible]

Query Match: [illegible]; Score [illegible]; DB: 2; Length: 459;
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy 1 MFAAVFMV[illegible] 60
Db 25 INAIVYSI[illegible] 84

Qy 61 [illegible] 120
Db 85 LKEILIKFF[illegible] 144

Qy 121 [illegible] 180
Db 145 LNAGIYL[illegible] 204

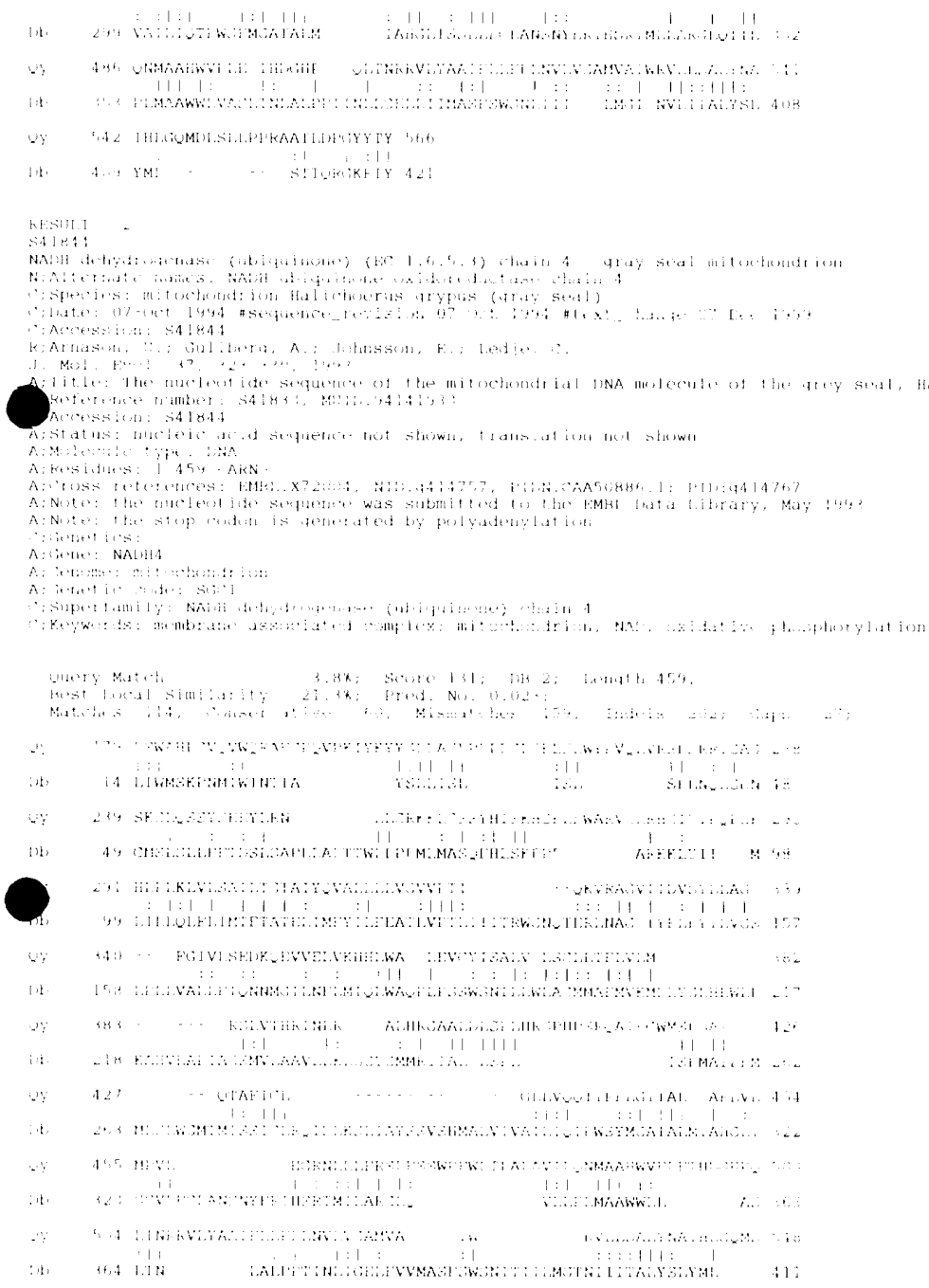
Qy 181 [illegible] 240
Db 205 LIIGLMN[illegible] 264

Qy 241 LLVGVVFL[illegible] 300
Db 265 [illegible] 324

Qy 301 [illegible] 360
Db 325 [illegible] 384

Qy 361 [illegible] 420
Db 385 [illegible] 444

Qy 421 [illegible] 459

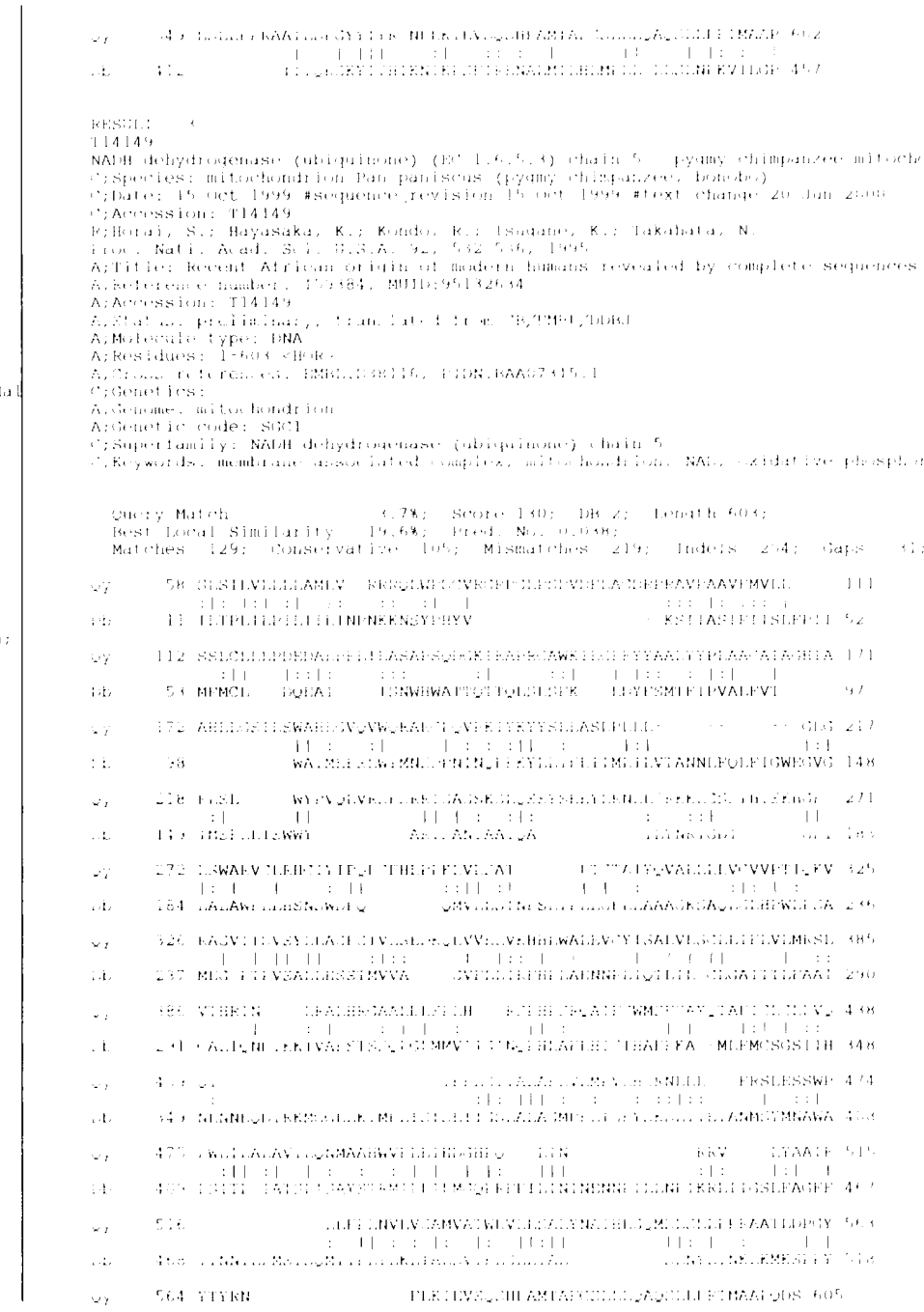


RESULT 2
S41844
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - gray seal mitochondrion
N;Alternate names: NADH ubiquinone oxidoreductase chain 4
C;Species: mitochondrion Halichoerus grypus (gray seal)
C;Date: 07-Oct 1994 #sequence_revision 07-Oct-1994 #text_change 07-Dec-1999
C;Accession: S41844
R;Arnason, U.; Gullberg, A.; Johnsson, E.; Ledje, C.
J. Mol. Evol. 42, 422-428, 1996
A;Title: The nucleotide sequence of the mitochondrial DNA molecule of the grey seal, Hal
A;Reference number: S41831; MUID:94341533
A;Accession: S41844
A;Status: nucleic acid sequence not shown, translation not shown
A;Molecule type: DNA
A;Residues: 1-459 <ARN>
A;Cross-references: EMBL:X72004; NID:g414757; PIDN:CAA50886.1; PID:g414767
A;Note: the nucleotide sequence was submitted to the EMBL data library, May 1992
A;Note: the stop codon is generated by polyadenylation
C;Genetics:
A;Gene: NADH4
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 3.8%; Score 131; DB 2; Length 459;
Best Local Similarity 21.3%; Pred. No. 0.02;
Matches 214; Conservative 70; Mismatches 179; Indels 202; Gaps 27;

RESULT 3
T14149
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 - pygmy chimpanzee mitochondrion
C;Species: mitochondrion Pan paniscus (pygmy chimpanzee, bonobo)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C;Accession: T14149
R;Horai, S.; Hayasaka, K.; Kondo, R.; Tsugane, K.; Takahata, N.
Proc. Natl. Acad. Sci. U.S.A. 92, 532-536, 1995
A;Title: Recent African origin of modern humans revealed by complete sequences of hom
A;Reference number: T14084; MUID:95132634
A;Accession: T14149
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-603 <HOR>
A;Cross-references: EMBL:D38116; PIDN:BAA07315.1
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylatio

Query Match 3.7%; Score 130; DB 2; Length 603;
Best Local Similarity 19.6%; Pred. No. 0.038;
Matches 129; Conservative 105; Mismatches 219; Indels 204; Gaps 31;



```
Db    519 IFYFSNM[illegible]FYFNIMHRS[illegible]PYIGLLISQNLPLLLLDLIWL   EKLLPKIISQYQVS 572
```

RESULT 4
QXMS4M
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - mouse mitochondrion
N;Alternate names: NADH ubiquinone oxidoreductase chain 4
C;Species: mitochondrion Mus musculus (house mouse)
C;Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 07-Dec-199[illegible]
C;Accession: A00440
R;Bibb, M.J.; Van Etten, R.A.; Wright, C.T.; Walberg, M.W.; Clayton, D.A.
Cell 26, 167-180, 1981
A;Title: Sequence and gene organization of mouse mitochondrial DNA.
A;Reference number: A00[illegible]; MUID:82[illegible]
A;Accession: A00440
A;Molecule type: DNA
A;Residues: 1-459 <BIB>
A;Cross-references: [illegible]; NID:[illegible]; PIDN:AAB48[illegible]; PID:[illegible]; EMBL:V00711; N[illegible]
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC[illegible]
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

```
  Query Match             [illegible]%;  Score 127.5;  DB 1;  Length 459;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible]43;
  Matches  112;  Conservative   82;  Mismatches  175;  Indels  219;  Gaps   27;

Qy     5[illegible] LSILVLLLAMLVRPFQIW[illegible]VRGK[illegible]PSFVL[illegible]LA[illegible]FAVFAAVFMV[illegible] 117
Db      6 LPSLMLLPLTWLSSPKKIWIN                    VISYSFLISLISLLLLW 43

Qy    118 LPDEDALPFLIIASAPSQ[illegible]AWA[illegible]YAA[illegible]AA AIA[illegible]H      1[illegible]
Db     44 QIDENYKN[illegible]SNMP[illegible]   [illegible]     LIIIIAWLLPIMLMAT[illegible]QHLLKKLNNVL 91

Qy    170 [illegible]IAAH[illegible]WAH[illegible]VLVWLSAETLVFEI    --------------SK 202
Db     92 QKIYISM[illegible]IILLMIPTAIEL[illegible]PFYILPFAILIPTL[illegible]KWGNQTHRLNAGIYFL 150

Qy    203 YYSLLASLPLLLGL---    GFLSLWYPVQLVFSFSRRTCACSKCL[illegible]YSEEYLRNLL 25[illegible]
Db    151 FYILI[illegible]SIPLLIALI[illegible]LIQNHV[illegible]LNL    MIL[illegible]     [illegible]LASWS    NNLL 1[illegible]

Qy    256 CRKKF[illegible]SYG[illegible]SKH[illegible]FLSWAPV[illegible]LRH[illegible]IY[illegible]HLPL KIVLCATLIGTAIYQVALLL 31[illegible]
Db    1[illegible]        W[illegible]MMAPL[illegible]LY[illegible]VHLWLPKAHVEA[illegible]MILSAILL 2[illegible]

Qy    315 LVC[illegible]VVP[illegible]KVPA[illegible]VI[illegible]  V[illegible]A[illegible]VI[illegible]LVVLLVSH[illegible]WA[illegible]AIVL 37[illegible]
Db    2[illegible]7 KLG    SYGMIRISIILLDPLIKYMAYPF ILLS       --LWG ----      268

Qy    37[illegible] [illegible]LIPIVIMS[illegible]VIHRINL[illegible]ALHK[illegible]AALLLS[illegible]LHRSPH[illegible]AIFWM[illegible]SAY QIAPI 431
Db    269 ---  MIMISSI[illegible]LRQIDLKS[illegible]--     ---- -   -  -----IAYSSVSHMALV 297

Qy    432 [illegible]LGLLV[illegible]IIFFI[illegible] ----   -IIALAPLVL   MPVLH[illegible]LLLFR[illegible]LESSWPPWLI 478
Db    298 IA[illegible]IML[illegible]PW[illegible]MDAIMLMIAH[illegible]LIL[illegible]LAN[illegible]HRIIM[illegible]AL[illegible]QMVFFLMAI 357

Qy    479 LALAVIL[illegible]NMAAHWVFLHIHL[illegible]H[illegible]QLINRRVLYAA        I[illegible]LNVLVGAMVA 529
Db    358 WWLMAS[illegible]ANLAI         PPSINLMGELFI[illegible]GLFSWSNF[illegible]IINI[illegible]GMYS 407

Qy    530 IWRVLLS   ALYNAI        HIAQMDLSLLPPRAAILDFGYYI 565
Db    408 MYMIIIIQRGKLINHMINLQPSHIRELLLMALHMIPLILLIISPKLII 455
```

RESULT 5
B58889
NADH dehydrogenase [illegible]
C;Species: mitochondri[illegible]
C;Date: 24-Oct-199[illegible] #sequence_revision 24-[illegible]
C;Accession: B5888[illegible]
R;Janke, A.; Gemme[illegible]
J. Mol. Evol. 42, [illegible]
A;Title: The mitoch[illegible]
A;Reference number: A5888[illegible]; MUID:97077[illegible]
A;Accession: B5888[illegible]
A;Status: nucleic [illegible]
A;Molecule type: DNA
A;Residues: 1-460 <JAN>
A;Cross-references: [illegible]
A;Note: submitted [illegible]/EMBL/DDBJ [illegible]
C;Genetics:
A;Map position: [illegible]
A;Genome: mitochondri[illegible]
A;Genetic code: SG[illegible]
C;Superfamily: NA[illegible]
C;Keywords: membrane-[illegible] complex; [illegible]

```
  Query Match        [illegible]
  Best Local Simil[illegible]
  Matches  104;  [illegible]  Mismat[illegible]

Qy    [illegible] VFAAVFMV[illegible]
Db    [illegible] INSSIYS[illegible]MLLINNENMSFSA[illegible]

Qy    [illegible] [illegible]YFLAAIAIA[illegible]AA[illegible]
Db    [illegible] MKLLMAHR[illegible]MAIAVFLIMYS[illegible]

Qy    [illegible] FQVFKIYF[illegible]WYFV[illegible]
Db    [illegible] LNAGIYFF[illegible]VALIFI[illegible]

Qy    2[illegible] [illegible]KHMFCWALV[illegible]
Db    1[illegible] NWGNLLL          WLA[illegible]
Qy    [illegible] [illegible]FVLAVVLLIVSYII[illegible]
Db    22[illegible] [illegible]SMVLAA[illegible]

Qy    [illegible] [illegible]VLMSFIVIHRINLSAL[illegible]
Db    2[illegible]
Qy    [illegible]
Db    [illegible]      W[illegible]ANLN[illegible]
Qy    4[illegible] [illegible]SWPPW[illegible]NMAAHWVFLHIHLHH[illegible]
Db    [illegible] M[illegible]SW[illegible]
Qy    52[illegible] LNVLVGAMVA[illegible]     LSALYNAI[illegible]
Db    [illegible] LNIVI[illegible]SHSISINFS[illegible]
```

RESULT 6
S26160
NADH dehydrogenase [illegible]
C;Species: mitoch[illegible]
C;Date: 31-Dec-199[illegible] #sequence_revision [illegible]
C;Accession: S261[illegible]
R;Arnason, [illegible]
J. Mol. Evol. 3[illegible]
A;Title: The complete [illegible] DNA [illegible]
A;Reference number: [illegible]; MUID:9227[illegible]

A;Accession: S261[illegible]
A;Molecule type: DNA
A;Residues: 1-459 <ARN>
A;Cross-references: EMBL:X63726; NID:g13431; PIDN:CAA45266.1; PID:g13441
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane associated complex; mitochondrion; NAD; oxidative phosphorylation;

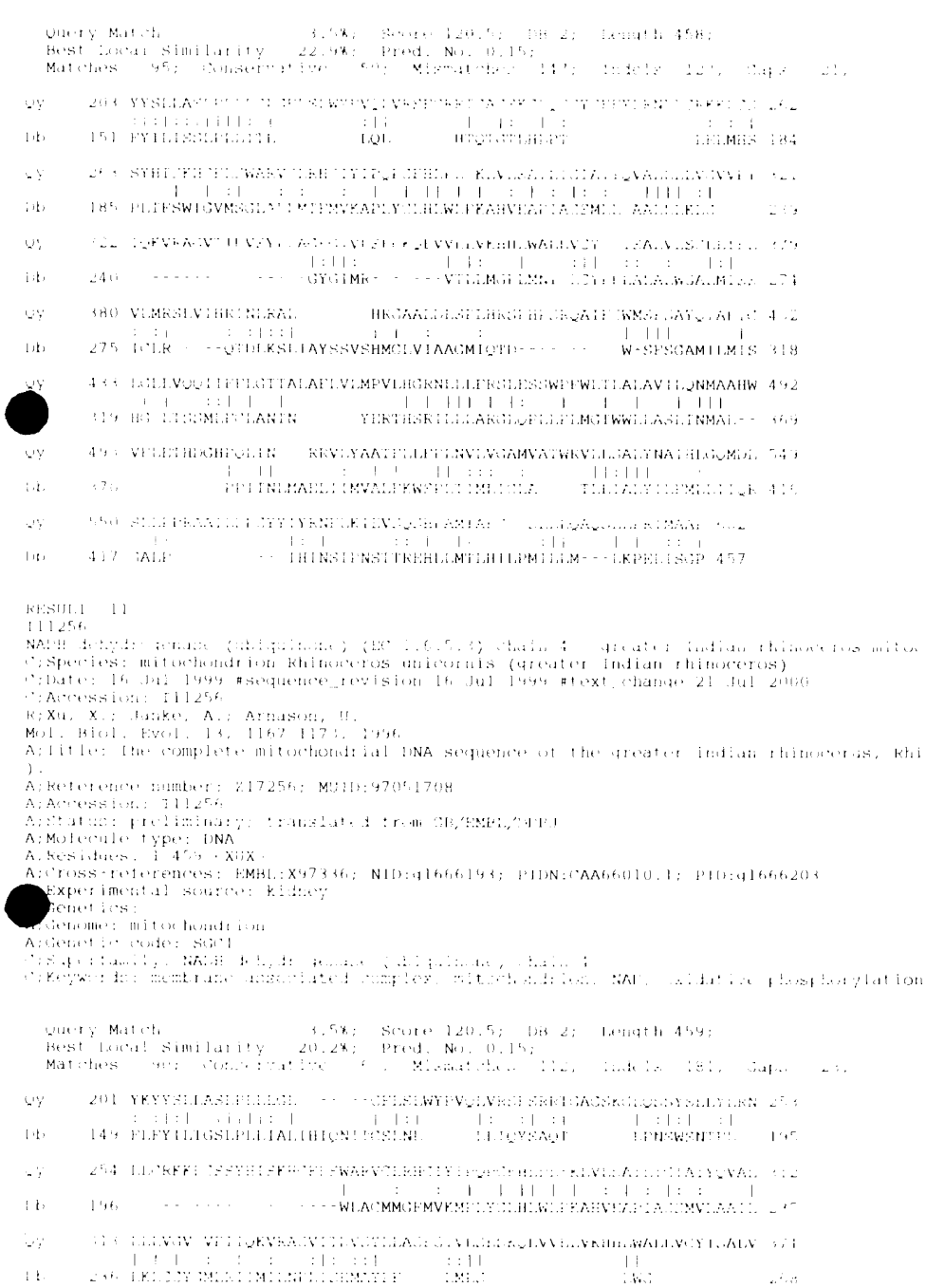
```
  Query Match              3.7%;  Score 127;  DB 2;  Length 459;
  Best Local Similarity   20.8%;  Pred. No. 0.047;
  Matches  107;  Conservative   66;  Mismatches  191;  Indels  [illegible];  Gaps   2[illegible];

Qy    175 LSWAHLGVQVWQKAL[illegible]PQVIF[illegible]YFYY[illegible]LLA[illegible] [illegible] WYPVQ[illegible]V[illegible] [illegible]A[illegible] 2[illegible]
Db     14 LIWMSKPNMIWINTIA----      YSLLISL     [illegible]      SP[illegible]NQL[illegible]N 48

Qy    2[illegible] [illegible]                LL[illegible]EKK[illegible]SIY[illegible]KHE[illegible]LSWARV[illegible]LRH[illegible]QIGP 290
Db     4[illegible] [illegible]WL[illegible]MLMAS[illegible]  [illegible]  M [illegible]

Qy    291 HLPLKLVLSATLTGTAIYQVALLLLVGVVPTI ---------QKVPACVTTDVSYLLAC- 339
Db     99 [illegible]LLQLF[illegible]IMIFIAIELIMF[illegible]EA[illegible]RWGNQ[illegible]RLNAG IYFL[illegible]VGS 157

Qy    340      P[illegible]IVL[illegible]EVVELVKHHLWA  [illegible]ALV [illegible]LVLMKSLV[illegible]H      388
Db    158 LPLLVALLFIQNNM[illegible]LNF[illegible]MIQ[illegible]WAQ[illegible]SWGN[illegible]A[illegible]MMA[illegible] [illegible]W[illegible] 2[illegible]

Qy    389 K[illegible]NLRALHRGA                 ALDE[illegible]LHES[illegible]HP[illegible]RQAI[illegible]W           42[illegible]
Db    218 KAHVEAP[illegible]VLAAVLLKL[illegible]MM[illegible]VL[illegible]MA[illegible]MLA[illegible]W[illegible]M[illegible] 2[illegible]

Qy    421      -MSFSAYQT--    AFICLGLLVQQIIFFLGTTALAFLVLMPVLHG KNLLLFRSLE 470
Db    278 RQIDLK[illegible]AYA[illegible]V[illegible]MA[illegible]IMA[illegible]LIQ[illegible]W[illegible]YMGA[illegible]A[illegible]     IA[illegible]AN [illegible]

Qy    471 SSW--      PFWLTLALAV[illegible]QNMAAHWVTLETHDGHPQLTNRRVLYAATFLLPPLNVLV 524
Db    332 SNYERTHSRIMILARGLQVL[illegible]PLMAAWWLL          ASLIN         LALPP[illegible]INL[illegible] 376

Qy    525 GAMVA       IW            RVLLGALYNAIHLGQMDLSLLPPRAATLLPGYYYYK N 568
Db    377 GELFVVMAS[illegible]SW[illegible]N[illegible]LM[illegible]N[illegible]IALYSLYML              [illegible]QRGN[illegible]HIK 424

Qy    569 FLKIEVGS[illegible]HPAMTAP[illegible]SLL[illegible]QAQSLLPRTMAAP 602
Db    425 NIKPSPIRLNALMILHLML[illegible] LLGLNPKIILGP 457
```

RESULT 7
I11309
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - smaller spotted catshark mitochon
C;Species: mitochondrion Scyliorhinus canicula (smaller spotted catshark, smaller spotte
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: I11309
R;Delarbre, C.; Spruyt, N.; Delmarre, C.; Gallut, C.; Barriel, V.; Janvier, P.; Laudet,
Genetics 150, 331-344, 1998
A;Title: The complete nucleotide sequence of the mitochondrial DNA of the Dogfish, Scyl
A;Reference number: Z17260; MUID:98393590
A;Accession: I11309
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-460 <DEL>
A;Cross-references: EMBL:Y16067; NID:g3618230; PIDN:CAA76028.1; PID:g3618240
A;Experimental source: muscle
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
A;Note: NADH 4
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane associated complex; mitochondrion; NAD; oxidative phosphorylatio

```
  Query Match              3.6%;  Score 125;  DB 2;  Length 460;
  Best Local Similarity   22.2%;  Pred. No. 0.067;
  Matches  112;  Conservative   66;  Mismatches  184;  Indels  142;  Gaps   24;

Qy     9[illegible] [illegible]LAG[illegible]KAV[illegible]AAV[illegible]MV[illegible]AL[illegible]AF[illegible]G[illegible]F[illegible]AWF[illegible]LGL 151
Db     56 YLAID PLSAP         LLLLPWLLP    -       LMILASQNHIIPEPLIRQRIYISL [illegible]

Qy    152 [illegible]AA[illegible]AA[illegible]AIAG[illegible]AAH[illegible]W[illegible]VW [illegible]LA[illegible]VP[illegible]YSLLASL 210
Db    100 L[illegible]SLQVFLIMAFSAIEMILFYIMFEA[illegible]LIPIL[illegible]RWGNQ[illegible]ERLNAGIYFLFYTLLGSL [illegible]

Qy    211 [illegible]               WY[illegible]V[illegible]VL[illegible]A[illegible]LENL[illegible]KLGSS 263
Db    160 PLLIALLPMQNDLNTLSMP       TIQYGHLPN      PSGWANKP---      ---  195

Qy    264 YFISKHG[illegible]LSWARV[illegible] KH[illegible]YI[illegible]KLVLSATL[illegible]GTAIYQVALLLVGV V[illegible] 320
Db    [illegible]               WW[illegible]A[illegible]AV[illegible]MM 245

Qy    321 TIQKVPACVTTDVSY-- [illegible]LACEGIVLSFDKQEVVELVKHHLWALEVCYISALVLSCLLT 377
Db    246 K[illegible]M[illegible]LA[illegible]GIVMI[illegible]            [illegible]LSQ[illegible]KGMIA 2[illegible]8

Qy    378 [illegible]LVLM[illegible]LV[illegible]ALHRGAALLL[illegible]LHRS[illegible]HL[illegible]RQAI[illegible]WMSFSAYQTAFICLGLLV 437
Db    [illegible]          M[illegible]NVAGAL[illegible]Q[illegible]          [illegible]AGAI[illegible]MIA[illegible] [illegible]

Qy    438 Q[illegible]IALA[illegible]LVLMPVLHGKNLLLFRSLLSSWLPWL[illegible]LALAVI[illegible]QNMA           489
Db    [illegible]21 [illegible]AL[illegible]ANI[illegible]         [illegible]A[illegible]MA[illegible]W[illegible]AN[illegible]AN[illegible] [illegible]7

Qy    490           AHWVTLETHDGHPQLTNRRVLYAATFLLPPLNVLVGAMVATWRVLLSA 537
Db    [illegible] [illegible]NWGNW[illegible]       [illegible]M[illegible]MN[illegible]N[illegible] 429

Qy    538 LYNAIHLGQMDLSLLPPRAATLDP 561
Db    430 GHIKEHL LLLLRVLLVLLLLKP 452
```

RESULT 8
S04756
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - rat mitochondrion
C;Species: mitochondrion Rattus norvegicus (Norway rat)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 07-Dec-1999
C;Accession: S04756; I51947; I65764; I65765
R;Gadaleta, G.; Pepe, G.; De Candia, G.; Quagliariello, C.; Sbisa, E.; Saccone, C.
J. Mol. Evol. 28, 497-516, 1989
A;Title: The complete nucleotide sequence of the Rattus norvegicus mitochondrial geno
A;Reference number: S04747; MUID:89362487
A;Accession: S04756
A;Molecule type: DNA
A;Residues: 1-459 <GAD>
A;Cross-references: EMBL:X14848; NID:g854269; PIDN:CAA32963.1; PID:g639984
R;Brown, G.G.; Castora, F.J.; Frantz, S.C.; Simpson, M.V.
Ann. N. Y. Acad. Sci. 361, 135-153, 1981
A;Title: Mitochondrial DNA polymorphism: evolutionary studies of the genus Rattus.
A;Reference number: I51947; MUID:81230588
A;Accession: I51947
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 308-363 <RES>
A;Cross-references: GB:M35251; NID:g[illegible]; PIDN:AAA[illegible].1; PID:g897634
R;Goddard, J.M.; Masters, J.N.; Jones, S.C.; Ashworth, W.D.
Chromosoma 82, 595-609, 1981
A;Title: Nucleotide sequence variants of Rattus norvegicus mitochondrial DNA.
A;Reference number: I52857; MUID:81259957
A;Accession: I65764
A;Status: preliminary

```
A;Molecule type: DNA
A;Residues: 
A;Cross-references: GB:M13696; NID:g343193; PIDN:AAA69848.1; PID:g897631
A;Accession: 
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 
A;Cross-references: GB:M13695; NID:g343199; PIDN:AAA69849.1; PID:g897633
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

  Query Match              5.5%;  Score 122.5;  DB 2;  Length 459;
  Best Local Similarity   19.1%;  Pred. No. 0.1;
  Matches  112;  Conservative   77;  Mismatches  177;  Indels  219;  Gaps   25;

Qy         62 LVLLLLAMLVKKRQLWHDVKGKPLPSPVDFLADDPPAVPAAVFMVLLSSLCLLLP-D 120
Db          9 IMLLPLIWLSANKKIWIN-                VISYSFLVSLLSLSLIWQND 46

Qy        121 EDALLFLLLASAPSQDGKIEAPRGAWKIDLEFYYAALYYPLAACATAGHIAAH------- 173
Db         47 ENY                                 LLLLLSWLLLMMLALNHHEKLNMMLEL 94

Qy        174 ---          LLGLLLSWAHLDVQVWQKAL            FKYYS 205
Db         95 YISMLISLQILLIMIFSATELF-ILFYILFEATLIPILLIIIRWGNQTERLNAGIYFLFYI 153

Qy        206 LLASLPLLLGH------GFLSLWYPVQLVKSFSEKICAGSKCLSSYSEEYLRNLLCRK 258
Db        154 LLGSIPLLIALISIQNSMGTLNF ----LILSLI-----  IHPIPSIWSNILL--  ---- 195

Qy        259 KLGSSYHISKHGFLSWARVCLRHFIYIFQFGFHLPL-KLVLSAILIGIATYQVALLLLVG 317
Db        196 ---     -------  -WLA MMAHLIEMPLYGVHLWLPKAHVEAPIAGSMILAAILLKLGG 240

Qy        318 VVPLLQKVWAGVLLP VSYLLAGPCIVLSHEDKQEVVELVKHHLKALEVCYISALVLSCLL 376
Db        241  -  YGMMRVSLLLLPLIKSLAYPFIILS---------- --LWG---    --------  268

Qy        377 IFLVLMRSLVIHRINLKALHPCAALGLSLLHRSFHHSRQAIPCVMSFSAY-QTAFICLGL 435
Db        269 --MIMISSICLRQIDLKSL-----  ---    --------       - IAYSSVSHMALVITAI 301

Qy        436 LVQQIFFFLI     ---   IIALAFLVL    MPVLHGRNLLLFRSLESSWPFWLLLALA 482
Db        302 MIQTPWSFMGAIMLMIAHGLISSLLFCLANINYERIHSRIMIMARGLQMIFPLMATWWLL 361

Qy        483 VLLNMAAHWVPLEHHLGHFLLIN    KKVLYAATP         LLFPLNVLVGAM       527
Db        362 ASLANLAL                  PPLINLMGELFIVMATFSWGNLLI LMAINIVITGHYSMYM 410

Qy        528  VAIWEVLLSALYNAI        HIGQMDLSLLPPAAILDPGYYI 565
Db        411 IIIIQRGKLISHMNNLQPSHIRELLIMALHIIPLMLLIINPKL I 455

RESULT   9
C90619
NADH dehydrogenase chain 5 ND5 (imported) - Apteryx haastii mitochondrion
C;Species: mitochondrion Apteryx haastii
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 03-Aug-2001
C;Accession: C90619
R;Haddrath, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A;Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogene
A;Reference number: A90613; MUID:21263106; PMID:11370967
A;Accession: C90619
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-5
A;Cross-references: 
C;Genetics:
A;Gene: ND5
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: NADH dehydrogenase (ubiquinone)
C;Keywords: mitochondrion

  Query Match              
  Best Local Similarity       Pred. No.
  Matches   96;  Conservative       Mismatches

Qy        154 YAALYYPLAACATAGHIAAHLLGSLLSWAHLDVQVWQ
Db            YSMLFLI                    WSLLL
Qy        212 LLL               WYPVQLVKSF
Db        133 LLIANNMH      WYI    LLFWWH
Qy        249 KYLRNLLCRK         LSWARV
Db        189 AWLASLLN          LAAA
Qy        303 LGIATYQVAL        KVFAGVLL        LVSY
Db        249 HSSIMVVA                NNL
Qy        358 HLWALEVCYISALV            IVLMRSLVIHRINLKA
Db        304                 ALLQLM        MVLL
Qy        418 FCWMSFSAY    
Db        349 L                  HSLAGFQLIEKM      K
Qy        457 VLHGRNLLLFRSLESS      WPFWLLLALAVLLNMAA
Db        389 GFYSKD                 NWALLLLLAISLAIYS
Qy        501 HPQLIN   KKVLYAATP   LLFPLNVL
Db        448 HPMVINL    AM    MA        SNILFIKIIEMMM
Qy        547 MDLSLLPPAAIL        
Db        508 MIHNLIF           NF 528

RESULT   10
B90623
NADH dehydrogenase chain 4 ND4 (imported) - Ap
C;Species: mitochondrion A   apteryx
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001
C;Accession: B90
R;Haddrath, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A;Title: Complete mitochondrial DNA genome sequences
A;Reference number: A90613; MUID:21263106; PMID:
A;Accession: B90
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-459
A;Cross-references: 
C;Genetics:
A;Gene: ND4
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: NADH dehydrogenase (ubiquinone)
C;Keywords: mitochondrion
```



Query Match 3.5%; Score 120.5; DB 2; Length 458;
Best Local Similarity 22.9%; Pred. No. 0.15;
Matches 95; Conservative 59; Mismatches 117; Indels 127; Gaps 21;

RESULT 11
I11256
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - greater Indian rhinoceros mitochondrion
C;Species: mitochondrion Rhinoceros unicornis (greater Indian rhinoceros)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: I11256
R;Xu, X.; Janke, A.; Arnason, U.
Mol. Biol. Evol. 13, 1167-1173, 1996
A;Title: The complete mitochondrial DNA sequence of the greater Indian rhinoceros, Rhino
).
A;Reference number: Z17256; MUID:97051708
A;Accession: I11256
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-459 <XUX>
A;Cross-references: EMBL:X97336; NID:g1666193; PIDN:CAA66010.1; PID:g1666203
A;Experimental source: kidney
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 3.5%; Score 120.5; DB 2; Length 459;
Best Local Similarity 20.2%; Pred. No. 0.15;
Matches 99; Conservative 64; Mismatches 112; Indels 181; Gaps 23;

RESULT 12
Q2B4M
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - bovine mitochondrion
N;Alternate names: NADH ubiquinone oxidoreductase chain 4
C;Species: mitochondrion Bos primigenius taurus (cattle)
C;Date: 28-Aug-1982 #sequence_revision 28-Aug-1982 #text_change 07-Dec-1999
C;Accession: A00439
R;Anderson, S., de Bruijn, M.H.L., Coulson, A.R., Eperon, I.C., Sanger, F. and Young, I
J. Mol. Biol. 156, 683-717, 1982
A;Title: Complete sequence of bovine mitochondrial DNA. Conserved features of the mam
A;Reference number: A00152; MUID:83010260
A;Accession: A00439
A;Molecule type: DNA
A;Residues: 1-459 <AND>
A;Cross-references: GB:J01394; NID:g13664; PIDN:AAB59227.1; PID:g529856; EMBL:V00654
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane associated complex; mitochondrion; NAD; oxidative phosphorylatio

Query Match 3.5%; Score 120; DB 1; Length 459;
Best Local Similarity 18.6%; Pred. No. 0.16;
Matches 76; Conservative 51; Mismatches 119; Indels 154; Gaps 15;

RESULT 13
I11543
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - spiny dogfish mitochondrion (frag
C;Species: mitochondrion Squalus acanthias (spiny dogfish)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: I11543
R;Rasmussen, A.S.; Arnason, U.
J. Mol. Evol. 48, 118-124, 1999
A;Title: Phylogenetic studies of complete mitochondrial DNA molecules place cartilaginou
A;Reference number: Z17282; MUID:99041711
A;Accession: I11543
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-460 <RAS>
A;Cross-references: EMBL:Y18141; NID:g4186605; PIDN:CAA77056.1; PID:g4186605
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC3
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 3.5%; Score 120; DB 2; Length 460;
Best Local Similarity 21.8%; Pred. No. 0.16;
Matches 121; Conservative 66; Mismatches 201; Indels 178; Gaps 27;

RESULT 14
I11629

NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - [illegible]
C;Species: mitochondrion [illegible] (A[illegible]
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: I11629
R;Johansen, S.; Bakke, [illegible]
Mol. Marine Biol. [illegible]
A;Title: The complete DNA sequence [illegible]
A;Reference number: [illegible] MUID:99041492
A;Accession: I11629
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-461 <JO>
A;Cross-references: [illegible]
A;Experimental source: [illegible]
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC3
A;Note: ND4
C;Superfamily: NADH dehydrogenase (ubiquinone)
C;Keywords: membrane [illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches [illegible]

RESULT 15
I10619
NADH dehydrogenase [illegible]
C;Species: mitochondrion [illegible]
C;Date: 20-Sep-1999 #sequence_revision [illegible]
C;Accession: I10619
R;Rasmussen, A.S.; [illegible]; Arnason, U.
J. Mol. Evol. 48, [illegible]
A;Title: The mit[illegible]
A;Reference number: [illegible]

```
A;Accession: T13819
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-458 <RAS>
A;Cross-references: EMBL:Y15189; NID:g3093351; PIDN:CAA75488.1; PID:g3093352
C;Genetics:
A;Genome: mitochondrion
A;Note: NADH4
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane-associated complex; [illegible] NAD; oxidative phosphorylation;

  Query Match              3.4%;  Score 117.5;  DB 2;  Length 458;
  Best Local Similarity   20.8%;  Pred. No. 0.25;
  Matches   77;  Conservative   55;  Mismatches  123;  Indels  115;  Gaps   15;

Qy        203 YYSLLASLPLLLGLGFLS   ---LWYPVQLVRSFSRRIGAGSKGLQSSYSEEYLRNLLCR 257
              ::|: |||||   | |||        |: |: :: |                    || |
Db        154 FYILIGSLPLLASLLFLSNILSTLYIPILMMES --      ------ ---  SEMY     -     190

Qy        258 FFL[illegible]WAFV[illegible]LEHF[illegible] 316
                             |   :  |              | ||  |||| |  |   : ||: |  ::    ||:| :
Db        191 -       -        HSPILWLGFFSALLIKTFLYGFHLWLLKAHVEATIAG-SMILAALMLKL 238

Qy        317 [illegible]VSTLLAGFGIVLSEDE[illegible] 376
              |                                    |:|::        :   ::   :   |     |  :  ||    :    ::
Db        239 G       -       -              GYGMI----RLSLMPFL--HSPKLSLILISIALWGAVM 271

Qy        377 IFLVLMRGLVTHRINLRALHRGAALDLGPLHRSPHPGRQAIF[illegible]WMSFSAYQTAF[illegible]L     433
              |    :  ::          :::|: |                                    :   :|         || |
Db        272 ISF[illegible]LP       -QIDLKAL    -          -          ---       IAYSSVSHMGLMIASIMILSL 306

Qy        434       GLLV[illegible]IIFFLGTFALAFLV   LMPVLHGRNLLLFRSLESSWF      FWLIILALAV 483
                     |    :     '       | ::|, :|                :  |  ::|  |    :      |       :|| : |
Db        307 WGMSGAFIMMIAHGLSSSALFFLANGNYEKTNTRTIILLRHTQMLLPLTGLWWLFILLTN 366

Qy        484 IIQNMAAHWVFLEIHDGHPQLINKKVLYAATFLLFFLNVLVGAMVATWRVLLSALYNAIH 543
              :       :  :::                  ::|        |:    :  |  ||      |       :     |:  :   :   |
Db        367 LAMPPSINFI             -SEIIIMSSLFIWSPLTFPFLALTMVITTTYSMSMFMLAQG K 417

Qy        544 LGQMDLSLLP 553
              |  :|     | |
Db        418 LSKMIKILSP 427


Search completed: August 31, 2002, 14:46:59
Job time: 146 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2002, 14:46:34 ; Search time 13.49 Seconds
(without alignments)
1914.451 Million cell updates/sec

Title: US-09-759-056-2
Perfect score: 3476
Sequence: 1 MSSQPAGNQTSPGATEDYSY..........NPHLQVFEKIALLGANGAQP 667

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt 40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	138	4.0	459	1	NU4M_RABIT	O79436 oryctolagus
2	[illegible]	[illegible]	459	1	NU4M_HALGR	P38601 halichoerus
3	[illegible]	[illegible]	603	1	NU5M_PANPA	P03916 pan paniscu
4	127.5	3.7	459	1	NU4M_MOUSE	P03911 mus musculu
5	127.5	3.7	460	1	NU4M_ORNAN	Q36458 ornithorhyn
6	[illegible]	[illegible]	459	1	NU4M_PHOVI	Q00506 phoca vitul
7	[illegible]	[illegible]	460	1	NU4M_SCYCA	O79410 scyliorhinu
8	122.5	3.5	459	1	NU4M_RAT	P05508 rattus norv
9	120.5	3.5	459	1	NU4M_CERSI	O03204 ceratotheri
10	120.5	[illegible]	459	1	NU4M_RHIUN	Q96068 rhinoceros
11	[illegible]	[illegible]	459	1	NU4M_BOVIN	P03910 bos taurus
12	[illegible]	[illegible]	460	1	NU4M_SQUAC	Q9ZZ45 squalus aca
13	[illegible]	3.4	460	1	NU4M_GADMO	P55781 gadus morhu
14	[illegible]	3.4	458	1	NU4M_MYXGL	Q9G2W9 myxine glut
15	[illegible]	3.4	459	1	NU4M_HIPAM	Q9ZZY2 hippopotamu
16	117.5	3.4	474	1	NU4M_DIDMA	P41308 didelphis m
17	117.5	3.4	[illegible]	1	NU5M_CHICK	P18940 gallus gall
18	[illegible]	3.4	460	1	NU4M_SALSA	Q9ZZM4 salmo salar
19	[illegible]	3.3	460	1	NU4M_ONCMY	P11631 oncorhynchu
20	115	3.3	603	1	NU5M_HUMAN	P03915 homo sapien
21	114.5	3.3	459	1	NU4M_DASNO	O21334 dasypus nov
22	114.5	3.3	460	1	NU4M_PIG	O79881 sus scrofa
23	114.5	3.3	461	1	NU4M_XENLA	P03912 xenopus lae
24	114	3.3	423	1	MIH5_DROME	Q9VQQ8 drosophila
25	114	3.3	459	1	NU4M_HORSE	P48655 equus cabal
26	114	3.3	459	1	NU4M_HYLLA	P03909 hylobates l
27	113	3.2	459	1	NU4M_CHICK	P18939 gallus gall
28	112.5	3.2	460	1	NU4M_CROLA	P34194 crossostoma
29	112.5	3.2	464	1	NU4M_PARLI	P12775 paracentrot
30	112	3.2	603	1	NU5M_PONPY	P03918 pongo pygma
31	111.5	3.2	459	1	NU4M_CANFA	Q95258 canis famil
32	111.5	3.2	460	1	NU4M_CARAU	O78687 carassius a
33	111	3.2	498	1	GUSA_ERWCH	Q47421 erwinia chr
34	111	[illegible]	[illegible]	[illegible]	NU4M_RHI[illegible]N	[illegible]
35	110	[illegible]	[illegible]	[illegible]	NU4M_SIR[illegible]A	[illegible]
36	110	[illegible]	[illegible]	[illegible]	N[illegible]_FAR[illegible]	[illegible]
37	110	[illegible]	[illegible]	[illegible]	NU[illegible]M_PAN[illegible]	[illegible]
38	109.5	[illegible]	[illegible]	[illegible]	NU4M_MA[illegible]	[illegible]
39	109	[illegible]	[illegible]	[illegible]	[illegible]_TYMV[illegible]	[illegible]
40	108.5	[illegible]	[illegible]	[illegible]	[illegible]X[illegible]_AS[illegible]	[illegible]
41	108	[illegible]	[illegible]	[illegible]	NU4M_BALM[illegible]	[illegible]
42	108	[illegible]	[illegible]	[illegible]	NU4M_SHEE[illegible]	[illegible]
43	108	[illegible]	[illegible]	[illegible]	NU4M_LAT[illegible]	[illegible]
44	108	[illegible]	[illegible]	[illegible]	[illegible]_TYMVA	[illegible]
45	107	[illegible]	[illegible]	[illegible]	NU4M_PEL[illegible]	[illegible]

ALIGNMENTS

RESULT 1
NU4M_RABIT
ID NU4M_RABIT STANDARD; PRT; 459 AA.
AC O79436;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4 (E[illegible]
GN MTND4 OR ND4 OR NADH4.
OS Oryctolagus cuniculus (Rabbit).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Ve[illegible]
OC Mammalia; Eutheria; Lagomorpha; Leporidae[illegible]
OX NCBI_TaxID [illegible]
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=983[illegible]
RA Gissi C., Gullberg A., Arnason U.;
RT "The complete [illegible] DNA sequence [illegible]
RT cuniculus.";
RL Genomics [illegible]
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone [illegible]
CC ----------
CC This SWISS-PROT entry is copyright. [illegible]
CC between the Swiss Institute of Bioinformatics [illegible]
CC the European Bioinformatics Institute. [illegible]
CC use by non-profit institutions as [illegible]
CC modified and this statement is not removed [illegible]
CC entities requires a license agreement [illegible]
CC or send an email to [illegible]
CC
DR EMBL; AJ001588; CAA[illegible]
DR InterPro; IPR[illegible]; NADH [illegible]
DR InterPro; IPR[illegible]
DR InterPro; IPR[illegible] N.
DR Pfam; PF00[illegible]
DR Pfam; PF0[illegible] N[illegible]
DR PRINTS; PR[illegible] N[illegible]XIDASE4.
KW Oxidoreductase; NAD; [illegible]; Mitochondri[illegible]
SQ SEQUENCE 459 AA; [illegible] MW; [illegible]

Query Match [illegible]; Score [illegible]
Best Local Similarity [illegible]; Pred. No. [illegible]
Matches [illegible]; Mismatches [illegible]

Qy 102 VPAAVFMV[illegible]A [illegible]

Db 2[illegible] INAIVYS[illegible] [illegible]

Qy 136 DGKIEAPE[illegible]AWE[illegible]YYAALYYPLAAPAIA[illegible] [illegible]

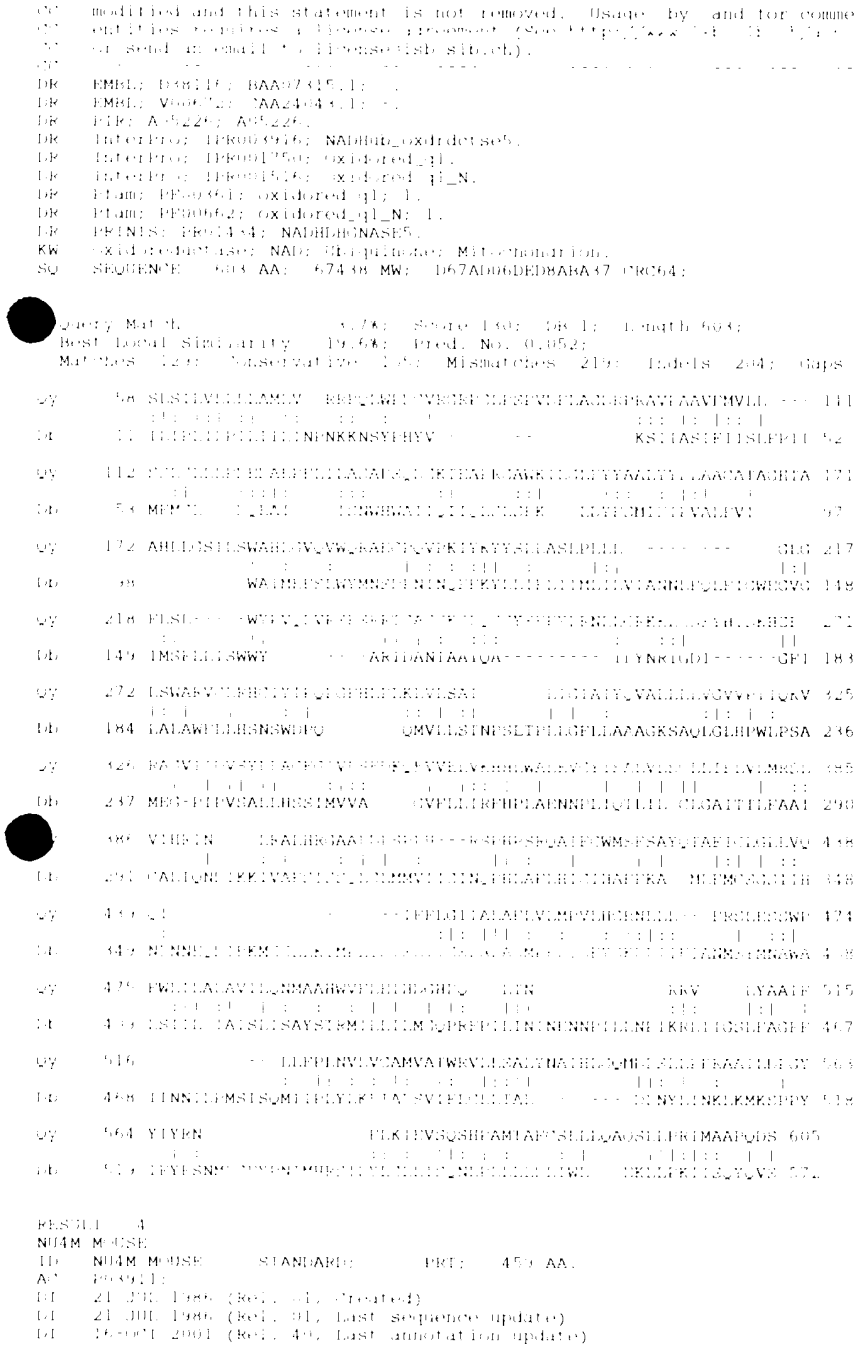
Db 8[illegible] FKEIL[illegible]VFLVMIFSA[illegible] [illegible]

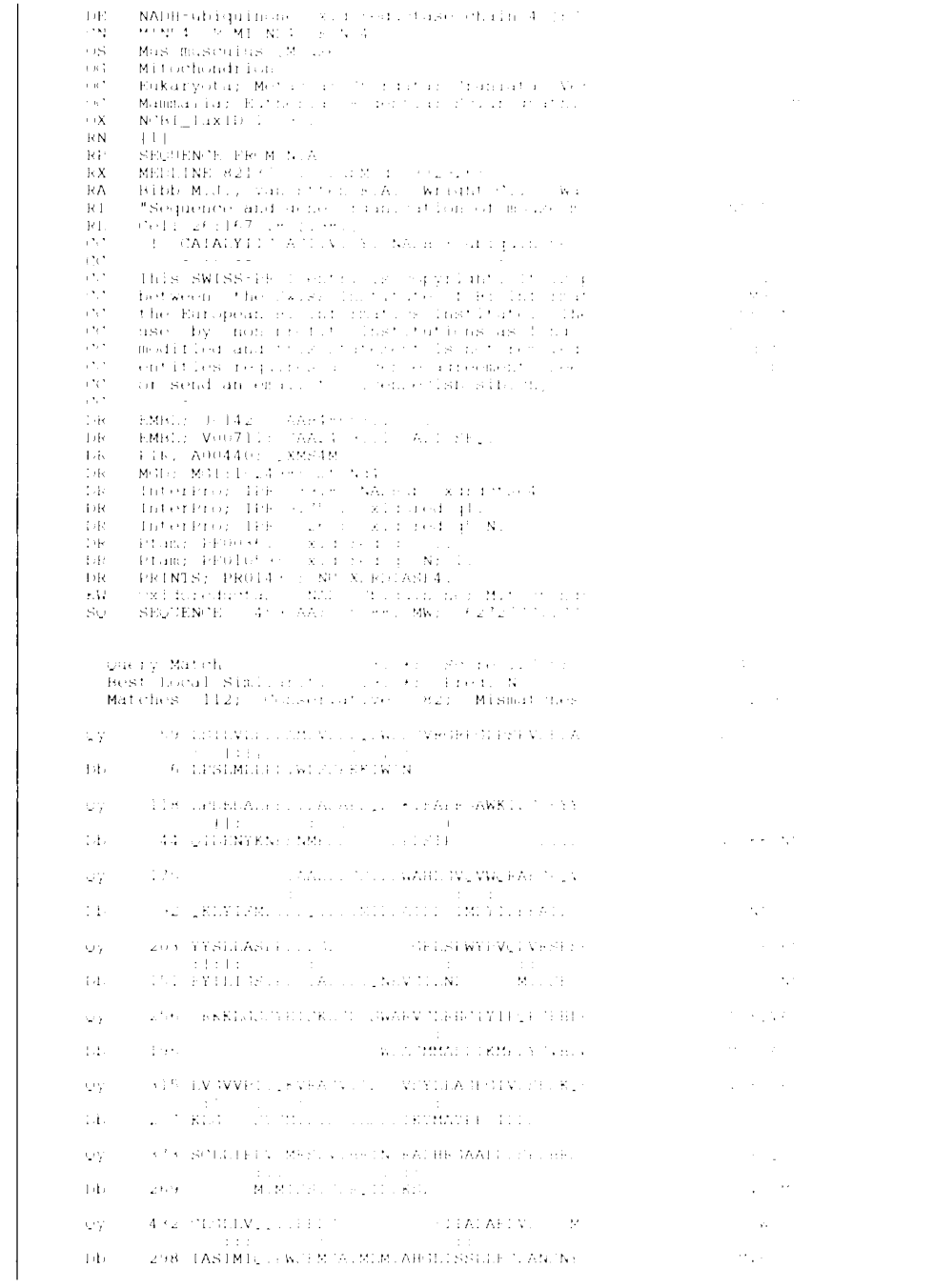
Qy 196 PQVPKIYY[illegible]WYPV[illegible] [illegible]

```
Db        143 LNA[illegible]YFL[illegible]Y[illegible]LPL[illegible]VALIYL          QN[illegible]M[illegible]        LNP 176

Qy        25[illegible] L[illegible]RFF[illegible]FH[illegible]WARV[illegible] [illegible]
Db        177 LLLQLMNK[illegible]MMAPLVKMPLY[illegible]LPKAHVEA[illegible]A[illegible]SMVLAA[illegible] 2[illegible]

Qy        314 LLVGVVF[illegible] [illegible]
Db        2[illegible]7 K[illegible]GMMK[illegible]YMAYF[illegible]             [illegible]     2[illegible]

Qy        374 [illegible]L[illegible]IFLVLMR[illegible]VIHRINLKALHRGAAL[illegible]HR[illegible]PHF[illegible]QA[illegible]FCWM[illegible]SAY [illegible]TAF[illegible] 4[illegible]2
Db        2[illegible]9        M[illegible]M[illegible]K[illegible]L                  [illegible]AY[illegible]SV[illegible]HMALV[illegible] 2[illegible]

Qy        4[illegible] [illegible]LLV[illegible]TALAFLVLMHVLL[illegible] [illegible]NLLLPR[illegible]W          [illegible]FWL[illegible]LALAVIL [illegible]5
Db        2[illegible] VA[illegible]LL[illegible]Q[illegible]PWSFMGA[illegible]ALM   [illegible] [illegible]AH[illegible]LLP[illegible]LAN[illegible]NY[illegible]R[illegible]MLLAR[illegible]Q[illegible] [illegible]2

Qy        [illegible] [illegible]NMAAHWVFLE[illegible]H[illegible]GHF    QL[illegible]NKRVLYAA[illegible]FLL[illegible]LNVLVGAMVA[illegible]WRVL[illegible]NA [illegible]41
Db        [illegible] PLMAAWKLVA[illegible]NLALF[illegible]NL[illegible]LL[illegible]MA[illegible]FGNWGNL[illegible]     L[illegible] [illegible]

Qy        542 IHLGQMDLSLLPPRAAILLPGYYIY 566
Db        4[illegible] YML[illegible]             [illegible]S[illegible]L[illegible]EFIY 4[illegible]
```

```
RESULT    2
NU4M_HALGR
ID   NU4M_HALGR     STANDARD;      PRT;   459 AA.
AC   P38[illegible];
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
GN   MTND4 OR ND4.
OS   Halichoerus grypus (Gray seal).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Halichoerus.
OX   NCBI_TaxID=9711;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94141933; PubMed=8308902;
RA   Arnason U., Gullberg A., Johansson E., Ledje C.;
RT   "The nucleotide sequence of the mitochondrial DNA molecule of the
RT   grey seal, Halichoerus grypus, and a comparison with mitochondrial
RT   sequences of other true seals.";
RL   J. Mol. Evol. 37:323-330(1993).
CC   -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X72004; CAA50886.1; -.
DR   PIR; S41844; S41844.
DR   InterPro; IPR003918; NADHbh_oxdrdtse4.
DR   InterPro; IPR001750; oxidored_q1.
DR   InterPro; IPR000260; oxidored_q5_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF01059; oxidored_q5_N; 1.
DR   PRINTS; PR01437; NUOXDRDTASE4.
KW   Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ   SEQUENCE   459 AA;  51641 MW;  D79A2515079049F6 CRC64;

  Query Match             3.8%;  Score 131;  DB 1;  Length 459;
  Best Local Similarity   21.3%;  Pred. No. 0.833;
  Matches  114;  Conservative   68;  Mismatches  159;  Indels  202;  Gaps   27;

Qy        1[illegible] [illegible]
Db         11 L[illegible]WMSKFNMIWINITA              Y[illegible]L[illegible]         [illegible]L          [illegible]LN[illegible]N 48

Qy        2[illegible] [illegible]IYIPQPGF 2[illegible]
Db         4[illegible] [illegible]M[illegible]MA[illegible]         A[illegible]KNLYI[illegible]    M [illegible]

Qy        2[illegible] HF[illegible]KLVL[illegible]AIL[illegible]AI[illegible]VA[illegible]V[illegible]VVF[illegible]          [illegible]KV[illegible]AGV[illegible]VSYLLAG  3[illegible]
Db         [illegible] [illegible]L[illegible]MIF[illegible]A[illegible]HLIMF[illegible]FEA[illegible]VF[illegible]W[illegible]NA[illegible]LFY[illegible] 1[illegible]7

Qy        3[illegible]       [illegible]G[illegible]VL[illegible]LDKQLVVELVKHIIGWA   LLV[illegible]LALV L[illegible]LLI[illegible]LVLM              3[illegible]2
Db        1[illegible]8 [illegible]LIVALLFIQNNM[illegible]LNPLM[illegible]MAS[illegible]WGNI[illegible]WLA[illegible]MMAPMVKML[illegible]LYGLHLWLP [illegible]7

Qy        3[illegible]3              K[illegible]V[illegible]HRINLK      ALHKGAA[illegible]HHSPHFSKQAIFCWMSFGAY  [illegible] 426
Db        2[illegible]8 KAHVL[illegible]A[illegible]AV[illegible]             [illegible]MATFM 2[illegible]2

Qy        427          QTAF[illegible]                                   [illegible]LLVQQ[illegible]FF[illegible]TAL   AFLVL 454
Db        2[illegible] M[illegible]WGM[illegible]M[illegible]  [illegible]MAL[illegible]VA[illegible]W[illegible]M[illegible]AH[illegible] [illegible]22

Qy        455 MPVL              HGRNLLLFRGLLGGWFFWLLLALAVI[illegible]NMAAHWVFLEIHDGHFQ 5[illegible]
Db        323 SGVLFCLANSNYERIHSRTMILARGLQ                   VLLPLMAAWWLL          AS 3[illegible]3

Qy        504 LINKKVLYAAIF[illegible]LF[illegible]LNVLVGAMVA        IW              [illegible]VLLSALYNAIHL[illegible]QMD 548
Db        364 LTN           LALFFFINL[illegible]GHLFVVMAGFGWGN[illegible]ILMGTNL[illegible]IALYGLYML      411

Qy        549 [illegible]RAA[illegible] NFLKILVSQSHFAMTAP[illegible]SLLLQA[illegible]SLLFKIMAAP 602
Db        412        [illegible]IQS[illegible]ENALMLLHLS[illegible]L[illegible]NFEVILGF 457
```

```
RESULT    3
NU5M_PANPA
ID   NU5M_PANPA     STANDARD;      PRT;   603 AA.
AC   P03916; Q35589;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
GN   MTND5 OR ND5.
OS   Pan paniscus (Pygmy chimpanzee) (Bonobo).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9597;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95132634; PubMed=7530363;
RA   Horai S., Hayasaka K., Kondo R., Tsugane K., Takahata N.;
RT   "Recent African origin of modern humans revealed by complete sequences
RT   of hominoid mitochondrial DNAs.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:532-536(1995).
RN   [2]
RP   SEQUENCE OF 1-79 FROM N.A.
RX   MEDLINE=82242101; PubMed=6284948;
RA   Brown W.M., Prager E.M., Wang A., Wilson A.C.;
RT   "Mitochondrial DNA sequences of primates: tempo and mode of
RT   evolution.";
RL   J. Mol. Evol. 18:225-239(1982).
CC   -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```





```
Qy     479 [illegible]LALAV[illegible]QNMAAHWV[illegible]NRRVLYAA [illegible]LNVLVCAHVA [illegible]
Db     356 [illegible]MASLANLAL [illegible] 407

Qy     530 IWRVLLS [illegible] [illegible]MA[illegible] 560
Db     408 [illegible]MAL[illegible] 45[illegible]
```

RESULT 5
NU4M_ORNAN
ID   NU4M_ORNAN   STANDARD;   PRT;   460 AA.
AC   Q36458;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
GN   MTND4 OR ND4.
OS   Ornithorhynchus anatinus (Duckbill platypus).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.
OX   NCBI_TaxID 9258;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Heart, and Liver;
RX   MEDLINE 97077000; PubMed 8919867;
RA   Janke A., Gemmell N., Feldmaier-Fuchs G., von Haeseler A.,
RA   Paabo S.;
RT   "The mitochondrial genome of a monotreme -- the platypus
RT   (Ornithorhynchus anatinus).";
RL   J. Mol. Evol. 42:153-159(1996).
CC   -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; X83427; CAA58431.1; -.
DR   InterPro; IPR003918; NADHdh_oxdrdtse4.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR000260; Oxidored_q5_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF01059; oxidored_q5_N; 1.
DR   PRINTS; PR01437; NUOXDRDTASE4.
KW   [illegible]
SQ   SEQUENCE   460 AA; 51740 MW; 5BD4F672CE739CA2 CRC64;

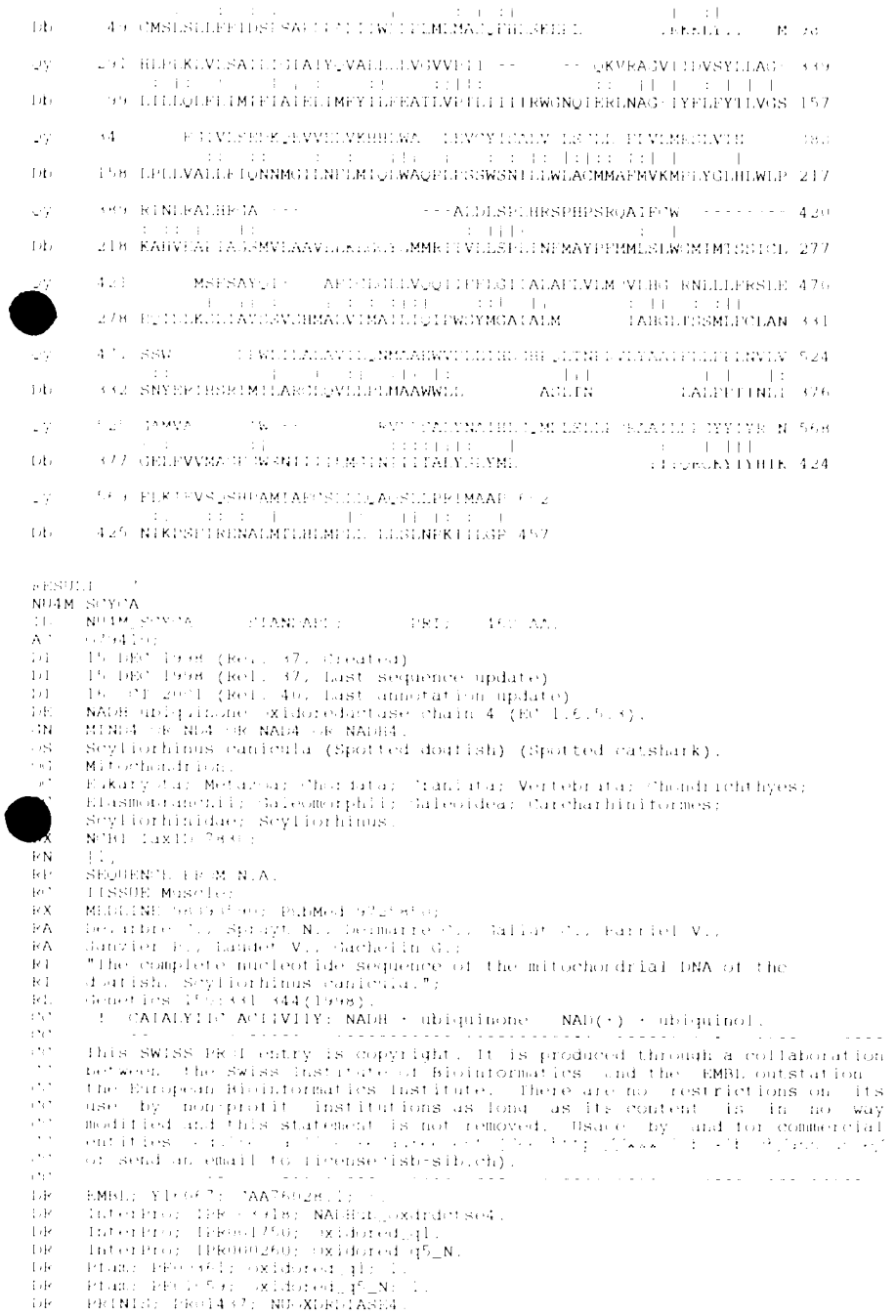
Query Match 3.7%; Score 127.5; DB 1; Length 460;
Best Local Similarity 19.1%; Pred. No. 0.058;
Matches [illegible]; Conservative 82; Mismatches 1[illegible]6; Indels 18[illegible]; Gaps 24;

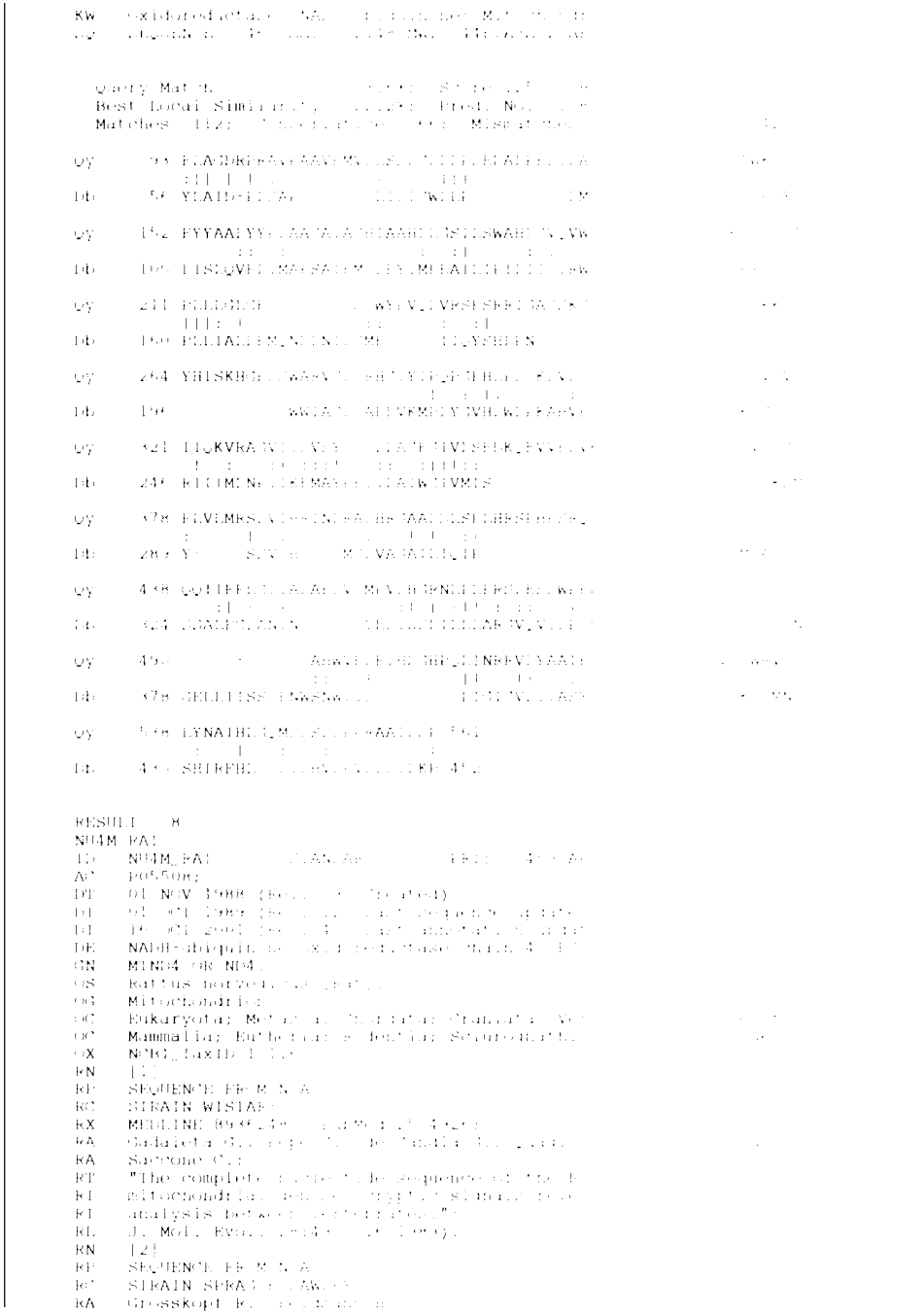
```
Qy     1[illegible] VPAAVPMVL[illegible] ALP[illegible]LAGAF[illegible]Q 135
Db      25 INSSIYSLIISSFSLLILNQHMD[illegible]LNPNMSFSA[illegible]SPLLILSPWLLPLMILASQPHL 84

Qy     136 DGK[illegible] 194
Db      85 MKFPMARKKTYLMLIVILQTALLMAFSAVELIMYYI-LFETTLIPTLIVIARWGNQTER 142

Qy     195 PQVFKIYKYYSLLASLPLL[illegible]SIWYPVQ[illegible]VRSFSPR[illegible]CAGS----K[illegible]QS 244
Db     143 LNAGLYFLFYTLLGSLPLLVALIFI-------QIQLGSLHILLLTLIPNPLQD 188

Qy     245 SY[illegible]LSWARV[illegible]KAVL[illegible]AL[illegible] 30[illegible]
Db     18[illegible] NWSNGIL WLACMLAFLVKMPLYG[illegible]HLWLPKAHVEAPIA 226

Qy     [illegible] [illegible] [illegible]
Db     227 [illegible]MVL AAILLKLG GYGIL RIILLLE [illegible] 2[illegible]

Qy     364 V[illegible]LMK[illegible] 41[illegible]
Db     2[illegible] [illegible]MAW[illegible]VVAA[illegible] [illegible]06

Qy     416 AIFCWMS[illegible]SAYQIA[illegible]LVLMPVLHGRNLLLFRSLE 470
Db     [illegible]07 W[illegible]MCA [illegible]ANLN [illegible]MVL[illegible] [illegible]4

Qy     471 SSWPFWLTLALAVILQNMAAHWVFLETHDGHP QLINRRVLYAAIF LLFP 519
Db     355 MSSW WLLASLANL [illegible] 397

Qy     520 LNVLVCAHVAIWRVL [illegible]ALYNATHL[illegible]MLL[illegible]FAAILDP 561
Db     398 LNTVIIAIYSTYMLITVQRG[illegible]INE[illegible]KGHM IMLLHLLPLTLLLLNP 451
```

RESULT 6
NU4M_PHOVI
ID   NU4M_PHOVI   STANDARD;   PRT;   459 AA.
AC   Q00506;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
GN   MTND4 OR ND4.
OS   Phoca vitulina (Harbor seal).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Phoca.
OX   NCBI_TaxID 9720;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 92277666; PubMed 1593642;
RA   Arnason U., Johnsson E.;
RT   "The complete mitochondrial DNA sequence of the harbor seal, Phoca
RT   vitulina.";
RL   J. Mol. Evol. 34:493-505(1992).
CC   -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; X63726; CAA45266.1; -.
DR   PIR; S26160; S26160.
DR   InterPro; IPR003918; NADHdh_oxdrdtse4.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR000260; Oxidored_q5_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF01059; oxidored_q5_N; 1.
DR   PRINTS; PR01437; NUOXDRDTASE4.
KW   Oxidoreductase, NAD, Ubiquinone, Mitochondrion.
SQ   [illegible]

Query Match 3.7%; Score 127; DB 1; Length 459;
Best Local Similarity 20.8%; Pred. No. [illegible];
Matches 107; Conservative 66; Mismatches 181; Indels 160; Gaps 25;

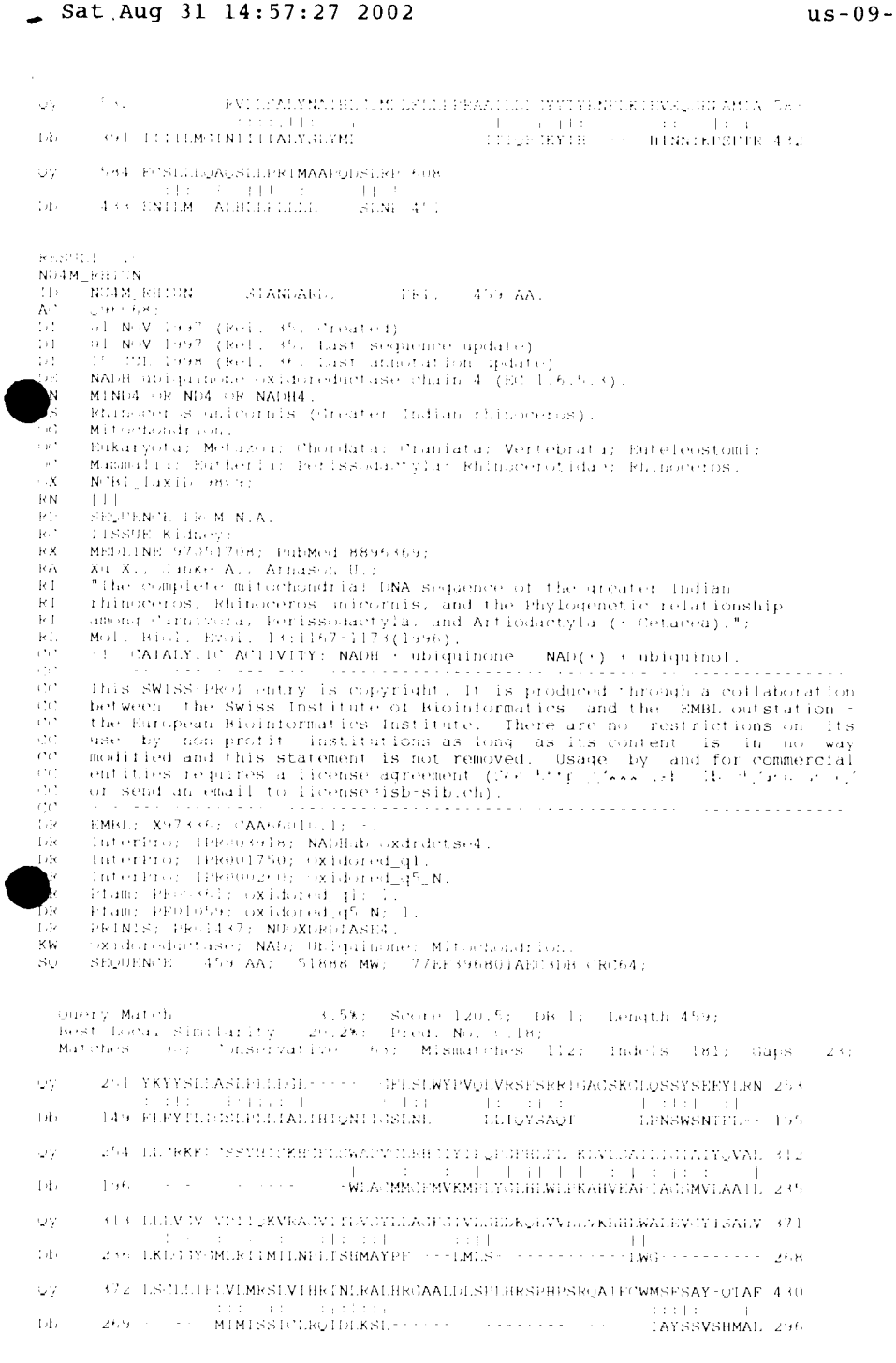
```
Qy     179 LSWAHLGVQVWQ[illegible] 238
Db      14 LIWMSKPNMIWINTTA YSLLISL -ISL SFLNQLGDN 48

Qy     239 SK[illegible]YLRN [illegible]LSWARV[illegible] 290
```

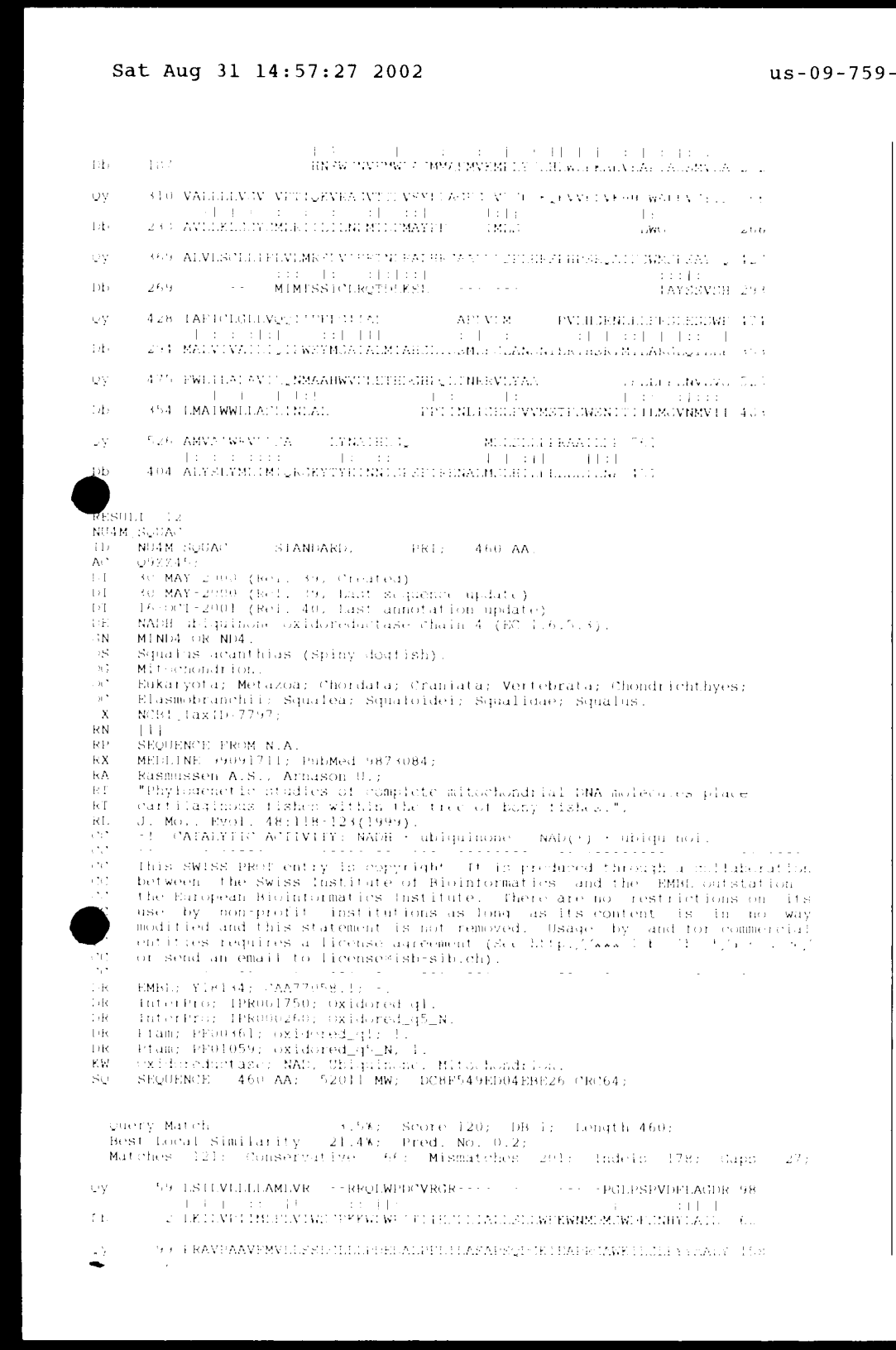




```
RT   "Analysis of a DNA segment from rat liver mitochondria containing the
RT   genes for the cytochrome oxidase subunits I, II, III, ATPase subunit
RT   6, and several tRNA genes.";
RL   Curr. Genet. 4:151-158(1981).
CC   -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X14848; CAA32963.1; -.
DR   PIR; S04756; S04756.
DR   InterPro; IPR003918; NADHub_oxdrdctse4.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR000260; Oxidored_q5_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF01059; oxidored_q5_N; 1.
DR   PRINTS; PR01437; NUOXDRDTASE4.
KW   Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
FT   CONFLICT    275    275       I -> V (IN REF. 2).
FT   CONFLICT    [illegible]    [illegible]       [illegible] -> [illegible] (IN REF. 2).
FT   CONFLICT    365    365       A -> S (IN REF. 2).
FT   CONFLICT    402    402       I -> T (IN REF. 2).
SQ   SEQUENCE   459 AA;  51800 MW;  8491C5501EE641FC CRC64;

  Query Match             3.5%;  Score 122.5;  DB 1;  Length 459;
  Best Local Similarity   19.1%;  Pred. No. 0.13;
  Matches  112;  Conservative   77;  Mismatches  177;  Indels  219;  Gaps   25;

Qy         62 LVLLLLAMLVKKRQLWFLVKGKIGLFSIVDFLAGDKIKAVIAAVIMVLLSGLVLLLL-V 120
Db          9 IMLLPLIWLSANKKIWTN                           VISYSFLVSLLSLSLLWQNL 16

Qy        121 EDALPFLLLLASAPSQDGKIFAPRGAWKIIGLFYYAALYYPLAACAIAGHIAAH------ 173
Db         47 ENYLNFGVMFSS       LFLSII           LIIILWAISNMSLASQSHMKKNMMHQKL 94

Qy        174 ----------LLGSILSWAHLCVQVWQPAFCPQVPFI---------------YKYYS 205
Db         95 YISMLISLQILLIMIFSATEL ILFYILFEAILIPTLIIITRWGNQTERLNAGIYFLFYI 153

Qy        206 LLASLPLLLGL            GFLCIWYPVQLVRSFSRKFCAIGKGLQSSYSEEYLENLLCKK 253
Db        154 LIGSIPLLIALISIQNSMGTLNF       LILSLT         THPLPSIWSNITL---  195

Qy        253 KLGSSYHISKHGFLSWARVCLRHCIYTPQPGFHLPL KLVLSATLGGIAIYQVALLLLVG 317
Db        196 --------- ---  -WLACMMAFMIKMPLYGVHLWLPKAHVEAPIAGSMILAAILLKLGG 240

Qy        318 VVLLIQKVRAGVLLLVSYLLAGLIVLSLHRQLVVLLVKHHLWALLVCLISALVLSCLL 376
Db        241     YGMMRVSIILDPLTKSLAYPFIILS ---------  LWG   --    -     -  268

Qy        377 IFLVLMKSLVTHRINLRALHRGAALDLSPLHRGFHPSRQAIPCWMSFSAY QIAFICLGL 435
Db        269   MIMISSICLRQIDLKSL                              IAYSSVSHMALVIIAI 301

Qy        436 LVQQIIFFLG--     ---   IIALAFLVL   -MPVLHGRNLLLFRSLESSWPFWLTLALA 482
Db        302 MILIPWGPMGAIMLMIAHGLISSALFCLANTNYERIHSRIMIMARGLQMIFPLMATWWLL 361

Qy        483 VILQNMAAHWVPLLITHDGHPQLIN    RKVLYAAIF          LLFPLNVLVGAM    -- 527
Db        362 ASLANLAL                 PPLINLMGELPIVMATFSWSNFSIILMAFNIVITGMYSMYM 410

Qy        528   VAIWKVLLSALTNAI          HLSQMDLSLLFIKAAILSGYYT 565
Db        411 IIITQRGKLISHMNNLQPSHTRELTLMALHIIPLMLLTINPKLIT 455
```

```
RESULT   9
NU4M_CERSI
ID   NU4M_CERSI     STANDARD;      PRT;   459 AA.
AC   O03204;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
GN   MTND4 OR ND4 OR NADH4.
OS   Ceratotherium simum (White rhinoceros) (Square-lipped rhinoceros).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.
OX   NCBI_TaxID 9807;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 97271644; PubMed 9126560;
RA   Xu X., Arnason U.;
RT   "The complete mitochondrial DNA sequence of the white rhinoceros,
RT   Ceratotherium simum, and comparison with the mtDNA sequence of the
RT   Indian rhinoceros, Rhinoceros unicornis.";
RL   Mol. Phylogenet. Evol. 7:189-194(1997).
CC   -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y07726; CAA69015.1; -.
DR   InterPro; IPR003918; NADHub_oxdrdctse4.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR000260; Oxidored_q5_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF01059; oxidored_q5_N; 1.
DR   PRINTS; PR01437; NUOXDRDTASE4.
KW   Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ   SEQUENCE   459 AA;  51719 MW;  F25C961D11EC5DFE CRC64;

  Query Match             3.5%;  Score 120.5;  DB 1;  Length 459;
  Best Local Similarity   20.2%;  Pred. No. 0.18;
  Matches   89;  Conservative   65;  Mismatches  112;  Indels  179;  Gaps   22;

Qy        201 YKYYSLLASLPLLLGL--   ---GFLSLWYPVQLVRSFSRKFCAGSKGLQSSYSEEYLRN 253
Db        149 FLFYILLIGSLPLLIALTHIQNLIGSLNL       LLIQYGIQI         LPSCWGNAFL   195

Qy        254 LLCKKKLGSSYHISKHGFLSWARVCLRHCIYTPQPGFHLPL KLVLSATLGGIAIYQVAL 312
Db        196 -  ------ ---          WLACMMAPMVKMPLYGLHLWLPKAHVEAPIAGSMVL AAI 234

Qy        313 LLLVGVVFLIQKVRAGVLLLVSYLLAGLIVLSLHRQLVVLLVKHHLWALLVCYISALVL 372
Db        235 LLKLG                           GYGMLR          IIMILNPL  --ISYMAYPFL 262

Qy        373 SCLLTPLVLMRSLVTHRINLRALHRGAALDLSPLHRGFHPSRQAIPCWMSFSAY QIAFI 431
Db        263 MLSLWGMIMTSSICLRQIDLKSL                                 IAYSSVSHMALV 297

Qy        432 CLGLLVQQIIFFLGTIAL  AFLVLMPVL       -         HGRNLLLFRSLESSWPFWLI 478
Db        298 IVAVLIQTPWSYMGAIALMIAHGLISSVLFCLANSNYERIHSRIMILARGLQTLLPL      354

Qy        479 LALAVILQNMAAHWVPLLITHDGHPQLINKKVLYAAIFLLLFPLNVLVGAMVA      TW     533
Db        355              MAMWWLL           ASLIN          LALPPTINLVGELFVVMSSFSWSN  3[illegible]
```





RESULT 12
NU4M_SQUAC
ID NU4M_SQUAC STANDARD; PRT; 460 AA.
AC Q92245;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NADH ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
GN MTND4 OR ND4.
OS Squalus acanthias (Spiny dogfish).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalea; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99091711; PubMed=9873084;
RA Rasmussen A.S., Arnason U.;
RT "Phylogenetic studies of complete mitochondrial DNA molecules place
RT cartilaginous fishes within the tree of bony fishes.";
RL J. Mol. Evol. 48:118-123(1999).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC ---------- --- ---- ---- --- -- --- -------- -- --------- -- ----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---------- --- ---- ---- --- -- --- -------- -- --------- -- ----
DR EMBL; Y16134; CAA77058.1; -.
DR InterPro; IPR001750; Oxidored_q1.
DR InterPro; IPR003260; Oxidored_q5_N.
DR Pfam; PF00361; Oxidored_q1; 1.
DR Pfam; PF01059; Oxidored_q5_N; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 460 AA; 52011 MW; DCBF549ED04EBE26 CRC64;

Query Match 5.5%; Score 120; DB 1; Length 460;
Best Local Similarity 21.4%; Pred. No. 0.2;
Matches 121; Conservative 66; Mismatches 201; Indels 178; Gaps 27;

RESULT 13
NU4M_GADMO
ID NU4M_GADMO STANDARD; PRT; 460 AA.
AC P55781;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NADH ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
GN MTND4 OR ND4.
OS Gadus morhua (Atlantic cod).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadoidei; Gadidae;
OC Gadus.
OX NCBI_TaxID=8049;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NORWEGIAN COASTAL 1;
RX MEDLINE=96411926; PubMed=9017926;
RA Johansen S., Bakke I.;
RT "The complete mitochondrial DNA sequence of Atlantic cod (Gadus
RT morhua): relevance to taxonomic studies among codfishes.";
RL Mol. Mar. Biol. Biotechnol. 5:203-214(1996).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC ---------- --- ---- ---- --- -- --- -------- -- --------- -- ----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---------- --- ---- ---- --- -- --- -------- -- --------- -- ----
DR EMBL; X99772; -; NOT_ANNOTATED_CDS.
DR InterPro; IPR000918; NADHUb_oxidoreds04.
DR InterPro; IPR001750; Oxidored_q1.

```
RP   SEQUENCE FROM N.A.
RA   Ursing B.M., Arnason U.;
RT   "Analyses of mitochondrial genomes strongly support a hippopotamus
RT   whale clade.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AJ010957; CAA09437.1; -.
DR   InterPro; IPR001750; oxidored_q1.
DR   InterPro; IPR000260; oxidored_q5_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF01059; oxidored_q5_N; 1.
KW   Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ   SEQUENCE   459 AA;  51915 MW;  5DC757C00D98DB11 CRC64;

  Query Match             3.4%;  Score 117.5;  DB 1;  Length 459;
  Best Local Similarity   19.2%;  Pred. No. 0.3;
  Matches   86;  Conservative   62;  Mismatches  116;  Indels  183;  Gaps   20;

Qy        201 YKYYSLLASLFLLL              GLGFLSLWTFVQLVKGFGKKIGAVGSKGLQSSISEEY 250
              : :::|:  | |::                 : :| | |   | |                 :|:|   :
Db        149 FLFYTLMGSLPLLVALIYIQNIFGSLNFLMLQYWTQAV                       GNSWGNVP 194

Qy        251 LRNLLCKKKLSSYHISKHCFLSWARVCLRHCIYTPQFCPHLFL KLVLSAILLGIAIYQ 309
              |                                  |   :    :   |   | || |  |   : |  : |: :
Db        195 L -----------------------WLACMMAFMVKMPLYGLHLWLPKAHVEAPIACSMVLA 232

Qy        310 VALLLLVGV----VPTIQKVRAGVTTDVSYLLAGFGIVLSEDKQEVVELVKHHLWALEVC 365
                 || | |      : ||              :: ::|    |:||                 ||
Db        233 AILLKLGGYGMLRITTI ------LNPLTEMMAYPFIMLS-------------LWG---- 268

Qy        366 YISAIVFSSIIFFLVLMESLVTHHTNLRALHRGAALFLSFLHRSPHPSRQAIPCWMSFSA 425
                                 :::  |:    :|: ::|                               :::|:
Db        269 -     -------  MIMTSSICLRQTDLKSL---------------  ----------IAYSS 290

Qy        426 Y-QIAFICLGLLVQQIIFFLCTIAL-        AFLVLM-----PVLHGRNLLLFRSLES 471
                  |  :  :  :|:|      ::| |||             : |    :           :| |  ::  | |::
Db        291 VSHMALVIVAILIQIPWSYMCATALMIAHGLISSMLFCLANSNYERIHSRTMILARGLQI 350

Qy        472 SWPFWLTLALAVILQNMAAHWVPLETHDGHPQLTNRRVLYAATFLLFPLNVLVGAMVA    529
                 |                 ||| ::                 |||              |  |      ||    ::
Db        351 LLPL---       -    --MAAWWLL  ---    -ASLIN-------  LALPPTINLVGELLVIM 383

Qy        530     IW             RVLLSALYNAIHLGQMDLSLLPFRAAILDPGYYIYR-NFLKIEVS 575
                    :|                :::: ||:     |                        |      | ||:   |  :       :
Db        384 SSFCWSNITILMGTNIMITTLYGLYML                FITQRGKYTHHINNITPSFT 431

Qy        576 QSHPAMTAFCSLLLQAQSLLPRTMAAP 602
              :  :  |: |      |  |     | |  :  :   |
Db        432 KEN ALMALHILPLLLLSLNPKIILGP 457


Search completed: August 31, 2002, 14:50:46
Job time: 252 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 31, 2002, 14:46:09 ; Search time 34.1 Seconds
(without alignments)
3383.801 Million cell updates/sec

Title: US-09-759-056-2
Perfect score: 3478
Sequence: 1 MSSQPAGNQTSPGATEDYSY.......NETLQVFKKTALLGANGAQP 667

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172004929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3478	100.0	667	4	Q9BX79	Q9bx79 homo sapien
2	3412.5	98.1	658	4	Q9BX78	Q9bx78 homo sapien
3	2903	83.5	560	4	Q9H9U8	Q9h9u8 homo sapien
4	2534.5	72.9	670	11	O70491	O70491 mus musculu
5	290.5	8.4	621	11	Q9DBN1	Q9dbn1 mus musculu
6	143.5	4.1	459	8	Q955T8	Q955t8 cheirogaleu
7	143.5	4.1	460	8	Q94SV2	Q94sv2 acrepelagadu
8	138.5	4.0	459	8	Q955T7	Q955t7 cheirogaleu
9	137	3.9	460	8	Q956R6	Q956r6 iguana igua
10	136	3.9	459	8	Q955U4	Q955u4 cheirogaleu
11	135	3.9	459	8	Q955U3	Q955u3 cheirogaleu
12	135	3.9	459	8	Q955T9	Q955t9 cheirogaleu
13	134.5	3.9	457	8	Q9BA05	Q9ba05 gonostoma g
14	134.5	3.9	457	8	Q955R7	Q955r7 daubentonia
15	133.5	3.8	459	8	Q9ME04	Q9me04 mus musculu
16	133	3.8	460	8	Q94SC0	Q94sc0 helicolenus
17	133	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
18	132.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
19	132.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
20	131	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
21	131	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
22	130.5	[illegible]	[illegible]	[illegible]	Q9GA02	[illegible]
23	130.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
24	130.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
25	130.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
26	130.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
27	130	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
28	130	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
29	130	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
30	129.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
31	129.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
32	129.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
33	129.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
34	129.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
35	129.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
36	129.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
37	129	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
38	129	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
39	128.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
40	128.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
41	128.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
42	128	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
43	127.5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
44	127	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
45	127	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

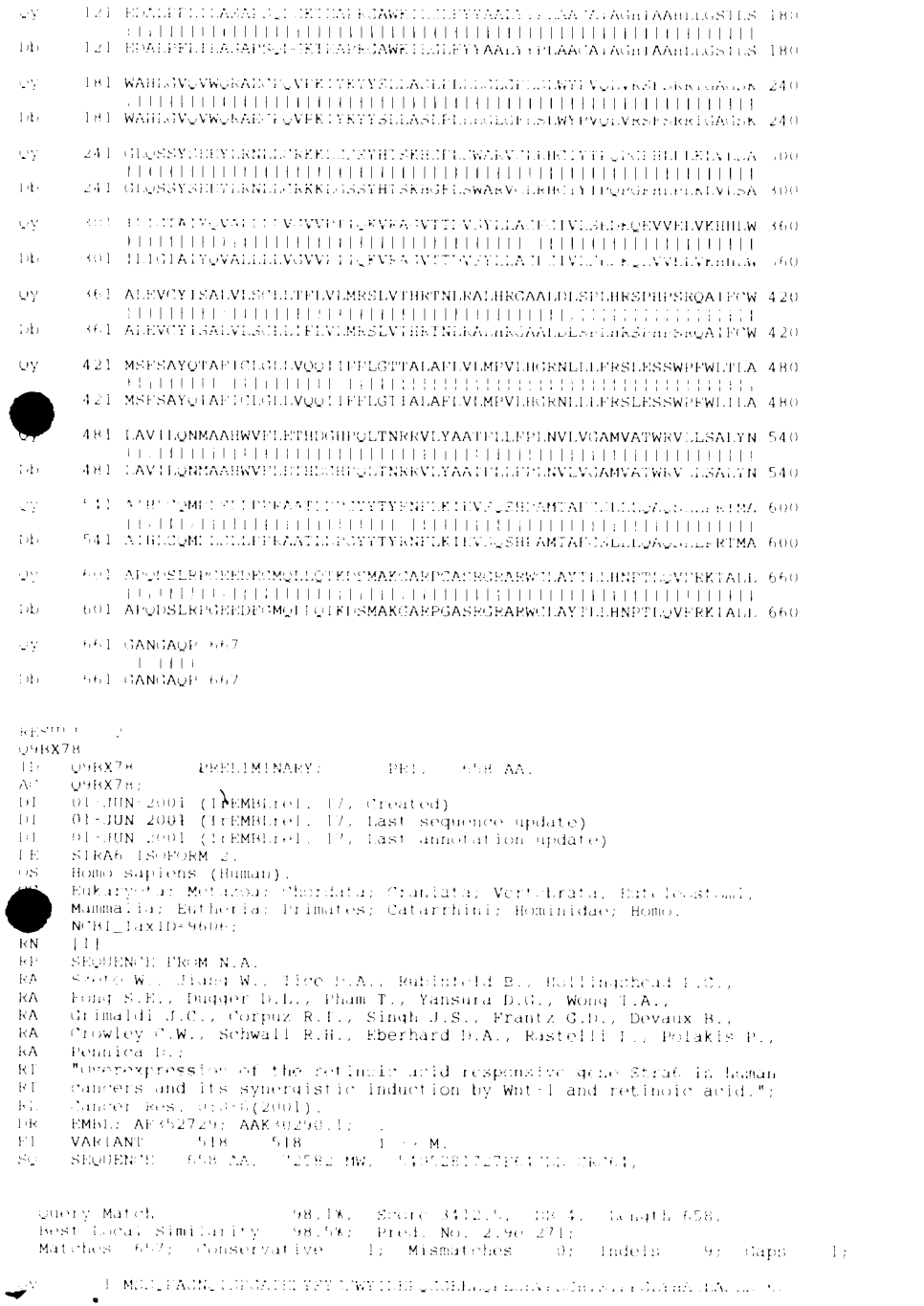
ALIGNMENTS

RESULT 1
Q9BX79
ID Q9BX79 PRELIMINARY; PRT; AA.
AC Q9BX79;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE STRA6 ISOFORM 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; ...
OC Mammalia; Eutheria; Primates; Catarrhini; ...
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Szeto W., ...
RA Fong S.E., ...
RA Grimaldi J.C., ...
RA Crowley C.W., ... Eberhard D.A., ...
RA Pennica D.;
RT "Overexpression ...
RT cancers and ...";
RL Cancer Res. ...
DR EMBL; AF352728; AAK...
FT VARIANT ...
SQ SEQUENCE ... AA; ... MW; ... CRC64;

Query Match 100.0%; ...
Best Local Similarity 100.0%; Pred. No. ...
Matches 667; ... Mismatches 0; ...

Qy 1 MSSQPAGN... 60
Db 1 MSSQPAGN... 60

Qy 61 LLVLLLLAMLVHRR... 120
Db 61 LLVLLLLAMLVHRR... 120



RESULT 7
Q9BX78
ID Q9BX78 PRELIMINARY; PRT; 658 AA.
AC Q9BX78;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SIRA6 ISOFORM 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Scott W., Hunt W., Tice D.A., Rubinfeld B., Rollinghead I.C.,
RA Ford S.E., Dugger D.L., Pham T., Yansura D.G., Wong T.A.,
RA Grimaldi J.C., Corpuz R.I., Singh J.S., Frantz G.D., Devaux B.,
RA Crowley C.W., Schwall R.H., Eberhard D.A., Rastelli I., Pelakis P.,
RA Pennica D.;
RT "Overexpression of the retinoic acid responsive gene Stra6 in human
RT cancers and its synergistic induction by Wnt-1 and retinoic acid.";
RL Cancer Res. 63:3-6(2001).
DR EMBL; AF352729; AAK30290.1; -.
FT VARIANT 518 518 I -> M.
SQ SEQUENCE 658 AA; 72582 MW; 5345281727F61714 CRC64;

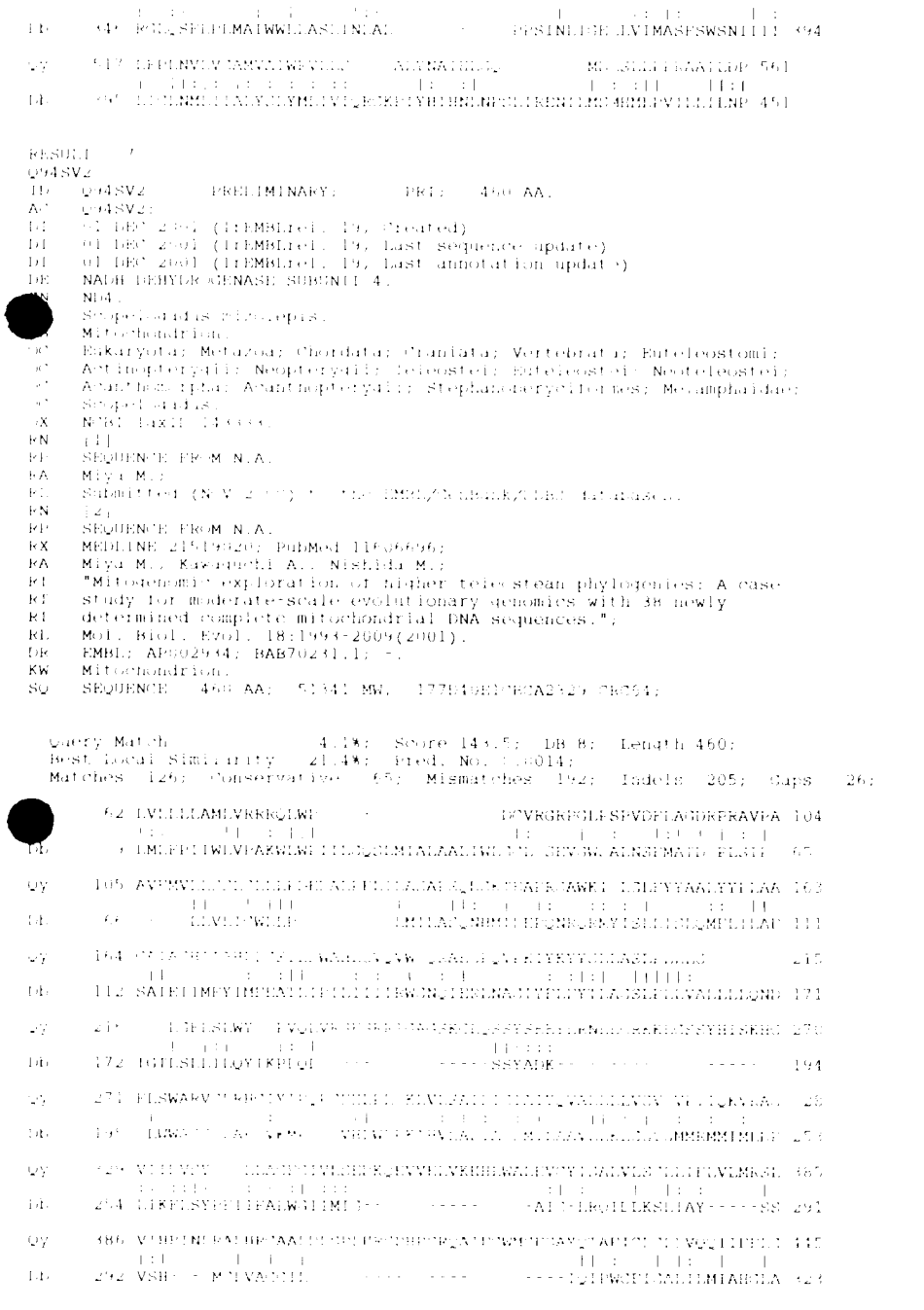
Query Match 98.1%; Score 3412.5; DB 4; Length 658;
Best Local Similarity 98.5%; Pred. No. 2.9e-271;
Matches 657; Conservative 1; Mismatches 0; Indels 9; Gaps 1;

RESULT 8
Q9H9U8
ID Q9H9U8 PRELIMINARY; PRT; 566 AA.
AC Q9H9U8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE MUSCULUS RETINOIC ACID-RESPONSIVE PROTEIN (SIRA6) MRNA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wadatsumi M., Hosoiri T., Kondo H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RL EMBL; AK0[illegible]; -.

```
SQ   SEQUENCE   560 AA;  61987 MW;  [illegible] CRC64;

  Query Match             83.5%;  Score 2903;  DB 4;  Length 560;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-229;
  Matches  560;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        108 MVLLSSPLLLLFDDDALFFLLLASAFSQDGK[illegible]AWE[illegible]YYAALYYPLAACATA 167
Db          1 MVLLSSPLLLLFDDDALFFLLLASAFSQDGK[illegible]YYAALYYPLAACATA 60

Qy        168 GHIAAHLLGSLLSWAHDLVLVWLRAESPLVFKIYKYYSLLASLP[illegible]FLSLWYPVQL 227
Db         61 GHIAAHLLGSLLSWAHDLVLVWLRAESPLVFKIYKYYSLLASLP[illegible]FLSLWYPVQL 120

Qy        228 V[illegible]WAKVLLRHTYTPQ 287
Db        121 V[illegible]WAKVLLRHTYTPQ 180

Qy        288 PGFHLPLKLVLSAILLGIAIYQVALLLLVGVVPIIQKVRAGVIIDVSYLLAGFGIVLSED 347
Db        181 PGFHLPLKLVLSAILLGIAIYQVALLLLVGVVPIIQKVRAGVIIDVSYLLAGFGIVLSED 240

Qy        348 KQEVVELVKHHLWALEV[illegible]LLLIFLVLMRGLVTHRKINLKALHKGAALDLSPLH 407
Db        241 KQEVVELVKHHLWALEV[illegible]LLLIFLVLMRGLVTHRKINLKALHKGAALDLSPLH 300

Qy        408 KSPHHSPQAIFCWMSFSAYQIAFI[illegible]TIALAFLVLMPVLHGRNLLLFR 467
Db        301 KSPHHSPQAIFCWMSFSAYQIAFI[illegible]TIALAFLVLMPVLHGRNLLLFR 360

Qy        468 SLESSWFEWLLLALAVILQNMAAHWVPLEIHDCHPQLINRRVLYAAIFLLFPLNVLVGAM 527
Db        361 SLESSWFEWLLLALAVILQNMAAHWVPLEIHDCHPQLINRRVLYAAIFLLFPLNVLVGAM 420

Qy        528 VAIWRVLLSALYNAIHLGQMDLSLLLFRAAIIL[illegible]YYIYRNELKIEVSQSHPAMIAFCSL 587
Db        421 VAIWRVLLSALYNAIHLGQMDLSLLLFRAAIIL[illegible]YYIYRNELKIEVSQSHPAMIAFCSL 480

Qy        588 LLQAQ[illegible]KIMAAPQDSL[illegible]GMQLLQTKDSMAKGARPCASRGRARWGLAYTLLH 647
Db        481 LLQAQ[illegible]KIMAAPQDSL[illegible]GMQLLQTKDSMAKGARPCASRGRARWGLAYTLLH 540

Qy        648 NPTLQVFRKTALLGANGAQP 667
Db        541 NPTLQVFRKTALLGANGAQP 560

RESULT 4
O70491
ID   O70491      PRELIMINARY;      PRT;   670 AA.
AC   O70491;
DT   1-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   RETINOIC ACID RESPONSIVE PROTEIN.
GN   STRA6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95377533; PubMed=7643975;
RA   Bouillet P., Oulad-Abdelghani M., Vicaire S., Garnier J.M.,
RA   Schuhbaur B., Dolle P., Chambon P.;
RT   "Efficient cloning of cDNAs of retinoic acid responsive genes in P19
RT   embryonal carcinoma cells and characterization of a novel mouse gene,
RT   Stra1 (mouse LERK 2/Eplg2).";
RL   Dev. Biol. 170:420-433(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97346723; PubMed=9203140;
RA   Bouillet P., [illegible] Mes[illegible]
RA   Chambon P.;
RT   "Developmental [illegible] pattern [illegible]
RT   gene encoding [illegible] membrane [illegible]
RL   Mech. Dev. [illegible]
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Bouillet P.;
RL   Submitted (A[illegible]) [illegible] EMBL/GenBank[illegible]
DR   EMBL; AF0624[illegible]
DR   MGD; MGI:107[illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible];  Score [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible]
  Matches  495;  [illegible]  Mismatches [illegible]

Qy          1 MSSQPA[illegible]SWYIDEPQ[illegible]
Db          1 MESQASEN[illegible]SWYIEEE[illegible]

Qy         59 LSILVLL[illegible]
Db         59 LSFLVLLL[illegible]

Qy        119 FDEDALFF[illegible]
Db        119 PDENPLPF[illegible]

Qy        179 LSWAHDLV[illegible]
Db        179 LSWAHFGV[illegible]

Qy        239 SKGLQSSY[illegible]
Db        239 SQGLQISY[illegible]

Qy        298 LSAILLG[illegible]
Db        299 ISAILLG[illegible]

Qy        358 HLWALEV[illegible]
Db        359 HLWIVEA[illegible]

Qy        418 FCWMSFSAY[illegible]
Db        419 VSWMSF[illegible]

Qy        478 LLALAVI[illegible]
Db        479 IVALAVI[illegible]

Qy        538 LYNAIHL[illegible]
Db        539 LYNIVHL[illegible]

Qy        598 IMAAP[illegible]
Db        599 PPLAPQ[illegible]

Qy        658 ALLGANGA[illegible]
Db        659 ALISAKANG[illegible]

RESULT 5
Q9DBN1
ID   Q9DBN1      PRELIMINARY;      PRT;   [illegible]
AC   Q9DBN1;
DT   01-JUN-2001 (TrEMBLrel. [illegible], Created)
DT   01-JUN-2001 (TrEMBLrel. [illegible], Last sequence [illegible])
```

```
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   1300002K09RIK PROTEIN.
GN   1300002K09RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=LIVER;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
DR   EMBL; AK005285; BAB23935.1; -.
DR   MGD; MGI:1921402; 1300002K09Rik.
SQ   SEQUENCE   621 AA;  71002 MW;  [illegible] CRC64;

  Query Match             8.4%;  Score 290.5;  DB 11;  Length 621;
  Best Local Similarity   21.7%;  Pred. No. 1.8e-15;
  Matches  141;  Conservative  115;  Mismatches  284;  Indels  111;  Gaps   20;

Qy         44 HISIPPGLYHACLASLSILVLLLLAMLVRR--RQLWPD- CVRGRPGLPSPVDFLAGDR 98
Db         20 HYSLIP---   ---SLCIILVLSFLQKREHRKQRKDDIGYLLQNHIHIIVFLDFVGIIG 75

Qy         99 PRAVPAAVFMVLLSSLCLLLPDEDALPFLTLASAPSQDGKTEAPRGAWK   ILGLFYYA 155
Db         76 NKWSYGAAFGAIARFVMFII [illegible] LQWALAFVIS [illegible] 121

Qy        156 ALYYPLAACATAT [illegible] 211
Db        122 LSYFPFFA [illegible] 181

Qy        215 GLGFLSLAT [illegible]
Db        182 [illegible] 231

Qy        271 FLSWAFV [illegible]
Db        235 EEJW [illegible] 281

Qy        3[illegible] [illegible]
Db        285 LD- [illegible]

Qy        377 IFLVLMRSLVIHKINKKALHRGAALLT [illegible] 4[illegible]
Db        334 [illegible]

Qy        437 VQQIIFFIGIIALAFLVLMFVLHGRNLLLFRSLESSWFFWLFLALAV   ILQNMAAHWV 493
Db        3[illegible] [illegible]

Qy        494 FLETHDGHPQ        LTNRRVLYAATFLLFPLNVLVGAMVATWRVLLSALYNAIHLGQM 547
```

```
Db        4[illegible] [illegible] 5[illegible]

Qy        548 [illegible] 5[illegible]
Db        5[illegible] [illegible] 553

Qy        [illegible] [illegible] 657
Db        5[illegible]   HEDKEIQKIVEYWCLN[illegible] 602


RESULT   6
Q95518
ID   Q95518      PRELIMINARY;      PRT;   459 AA.
AC   Q95518;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NADH DEHYDROGENASE SUBUNIT 4.
GN   ND4.
OS   Cheirogaleus major (Lemur).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
OC   Cheirogaleus.
OX   NCBI_TaxID=47177;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JP138;
RX   MEDLINE=21184272; PubMed=11286490;
RA   Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
RT   "Molecular phylogeny of the lemur family Cheirogaleidae (primates)
RT   based on mitochondrial DNA sequences.";
RL   Mol. Phylogenet. Evol. 19:45-56(2001).
DR   EMBL; AF224618; AAK70526.1; -.
KW   Mitochondrion.
SQ   SEQUENCE   459 AA;  51807 MW;  D00DCDEFF06CD6AA CRC64;

  Query Match             4.1%;  Score 143.5;  DB 8;  Length 459;
  Best Local Similarity   18.8%;  Pred. No. 0.0014;
  Matches  101;  Conservative   98;  Mismatches  183;  Indels  153;  Gaps   19;

Qy         90 FVDFLAGDRPRAVPAAVFMVLLSSLCLLLPDED                        A 123
Db         1[illegible] [illegible] 72

Qy        124 [illegible] 1[illegible]
Db         73 [illegible] 130

Qy        1[illegible] [illegible] 242
Db        131 [illegible]   VAL[illegible]LLNMAT-- 178

Qy        2[illegible] [illegible] KLVLSAI [illegible]
Db        179         [illegible] 224

Qy        3[illegible] [illegible] VKHHLW [illegible]
Db        225 [illegible]M.KTLAY                            259

Qy        3[illegible] ALEV [illegible] 420
Db        260        [illegible]RQTDLKSL          ---    -          285

Qy        421 MGFSAY QTAFICLGLLVQQIIFFIGIIAL          AFLVLM    --FVLHGRNLLLF 466
Db        286 [illegible] 3[illegible]

Qy        467 RSLESSWFFWLFLALAVILQNMAAHWVFLETHDGHPQLTN     RRVLYAA      TFL 516
```

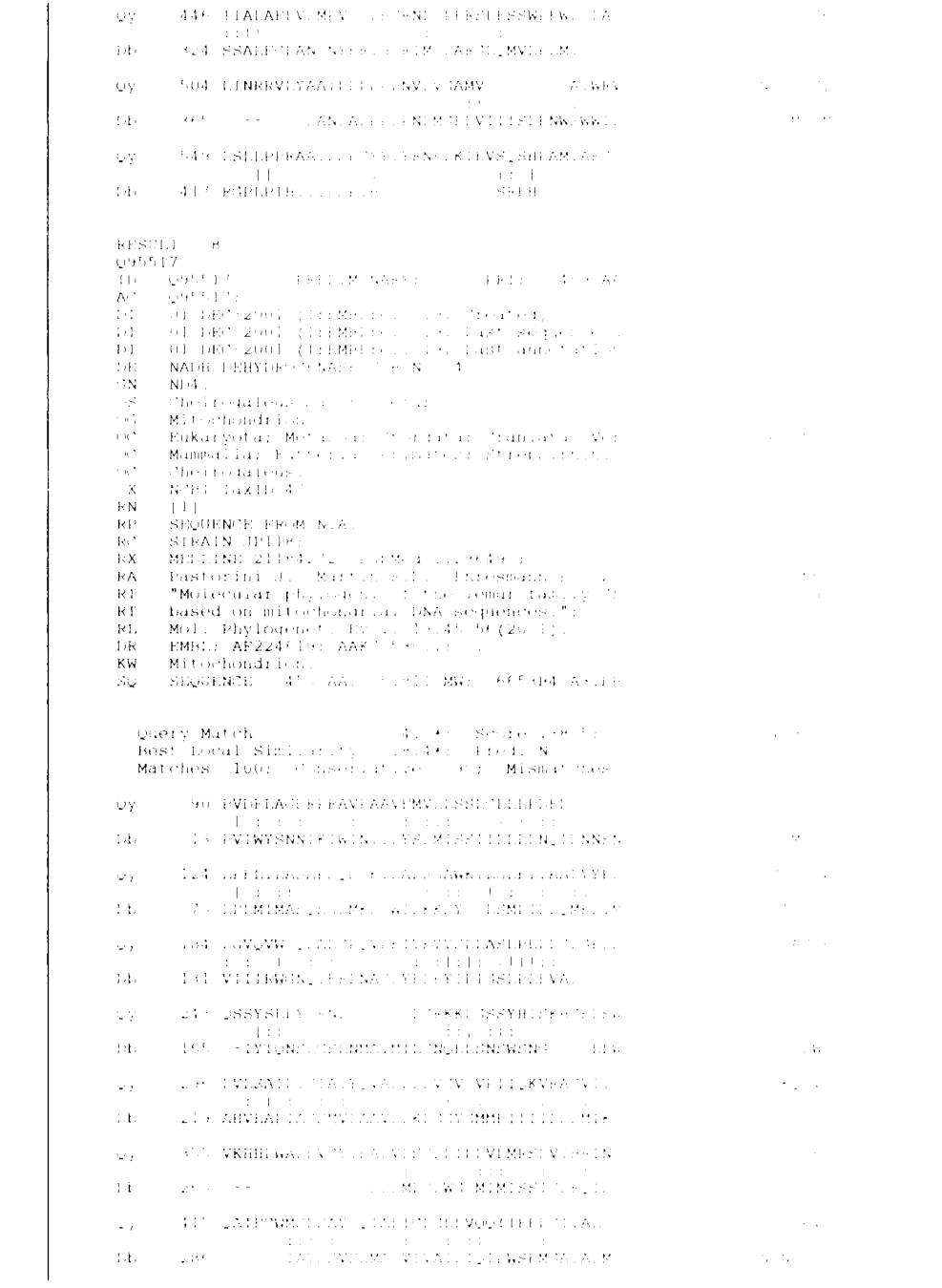


RESULT 7
Q94SV2
ID Q94SV2 PRELIMINARY; PRT; 460 AA.
AC Q94SV2;
DT 02-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 4.
GN ND4.
OS Scopelosaurus [illegible].
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Aulopiformes; Aulopoidei; Stephanoberycoidei; Melamphaidae;
OC Scopelosaurus.
OX NCBI_TaxID=[illegible];
RN [1]
RP SEQUENCE FROM N.A.
RA Miya M.;
RL Submitted (NOV-2001) to EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21519620; PubMed=11679696;
RA Miya M., Kawaguchi A., Nishida M.;
RT "Mitogenomic exploration of higher teleostean phylogenies: A case
RT study for moderate-scale evolutionary genomics with 38 newly
RT determined complete mitochondrial DNA sequences.";
RL Mol. Biol. Evol. 18:1993-2009(2001).
DR EMBL; AP002994; BAB70231.1; -.
KW Mitochondrion.
SQ SEQUENCE 460 AA; 51341 MW; [illegible] CRC32;

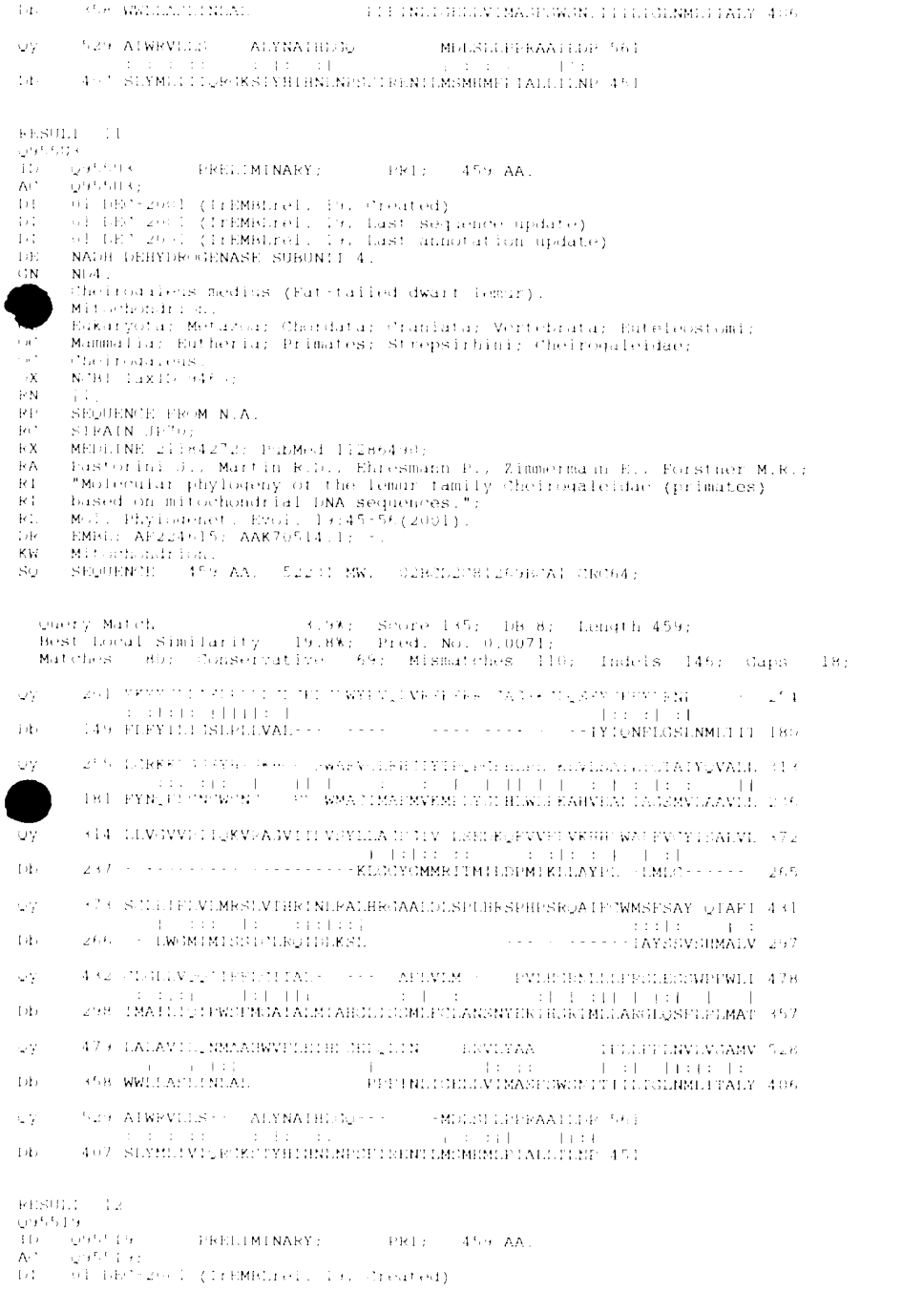
Query Match 4.1%; Score 143.5; DB 8; Length 460;
Best Local Similarity 21.4%; Pred. No. [illegible];
Matches 126; Conservative 65; Mismatches 192; Indels 205; Gaps 26;

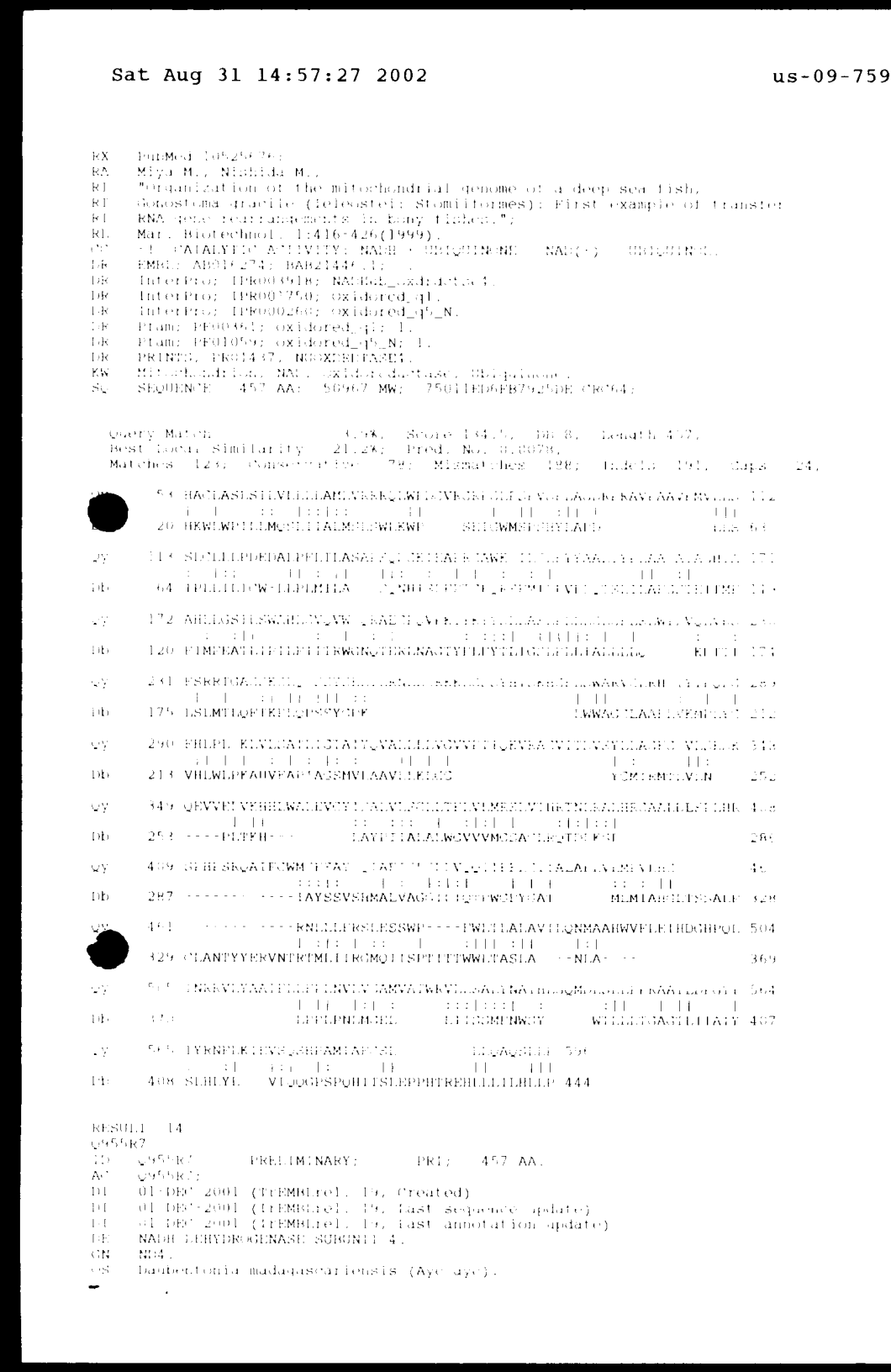
RESULT 8
Q95917
ID Q95917 [illegible]
AC Q95917;
DT 01-DEC-2001 (TrEMBLrel. [illegible]
DT 01-DEC-2001 (TrEMBLrel. [illegible]
DT 01-DEC-2001 (TrEMBLrel. [illegible]
DE NADH DEHYDROGENASE [illegible] 4
GN ND4.
OS Chaetodontoplus [illegible]
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Ver[illegible]
OC Mammalia; Eutheria; [illegible]
OC Chaetodontoplus.
OX NCBI_TaxID=4[illegible]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JPTIP[illegible]
RX MEDLINE=21[illegible]
RA Pastorini J., Mar[illegible]
RT "Molecular phylogenetics [illegible]
RT based on mitochondrial DNA sequences.";
RL Mol. Phylogenet. [illegible]
DR EMBL; AF224[illegible]
KW Mitochondrion.
SQ SEQUENCE [illegible]

Query Match [illegible]
Best Local Similarity [illegible]
Matches 100; [illegible]



```
Db    449 LNE 451

RESULT 9
Q956R6
ID   Q956R6       PRELIMINARY;      PRT;   460 AA.
AC   Q956R6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NADH DEHYDROGENASE SUBUNIT 4.
GN   NAD4.
OS   Iguana iguana (Common iguana).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Iguania; Iguanidae; Iguaninae; Iguana.
OX   NCBI_TaxID=8517;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21190842; PubMed=11297180;
RA   Janke A., Erpenbeck D., Nilsson M., Arnason U.;
RT   "The mitochondrial genomes of the iguana (Iguana iguana) and the
RT   caiman (Caiman crocodylus): implications for amniote phylogeny.";
RL   Proc. R. Soc. Lond., B, Biol. Sci. 268:623-631(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Janke A.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ278511; CAC37092.1; -.
KW   Mitochondrion.
SQ   SEQUENCE   460 AA;  51509 MW;  6A087DF61FF78F83 CRC64;

  Query Match             [illegible]%;  Score 137;  DB 8;  Length 460;
  Best Local Similarity   23.9%;  Pred. No. 0.0049;
  Matches  128;  Conservative   80;  Mismatches  187;  Indels  218;  Gaps   31;

Qy     41 PSCHISIPPGLYHACLASLSILVLLLLAMLVRKRQLWPDCVRGRPGLLSEVDF     LAT 96
Db     13 PTILTTKAPHINIPTAYSILLALLCTWL              KPTLTTNTTTTNLLLAV 59
[illegible]
Db    164 QIFLLLIFSSINLILFYIMPEATLIFTLLLTKWGNQAEKLNAGT                146
[illegible]
Db    149        YFLFYILASSLPLLIAILYLN                    IENTHLSMILL 180
[illegible]

RESULT 10
Q95504
ID   Q95504       PRELIMINARY;      PRT;   459 AA.
AC   Q95504;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   NADH DEHYDROGENASE SUBUNIT 4.
GN   ND4.
OS   Cheirogaleus medius (Fat-tailed dwarf lemur).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
OC   Cheirogaleus.
OX   NCBI_TaxID=9460;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JP6;
RX   MEDLINE=21184272; PubMed=11286490;
RA   Pastorini J., Martin R.D., Ehresmann P., Zimmermann E., Forstner M.R.;
RT   "Molecular phylogeny of the lemur family Cheirogaleidae (primates)
RT   based on mitochondrial DNA sequences.";
RL   Mol. Phylogenet. Evol. 19:45-56(2001).
DR   EMBL; AF224614; AAK70510.1; -.
KW   Mitochondrion.
SQ   SEQUENCE   459 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score 13[illegible];  DB 8;  Length 459;
  Best Local Similarity   19.8%;  Pred. No. 0.0059;
  Matches   89;  Conservative   66;  Mismatches  113;  Indels  146;  Gaps   17;

Db    149 FLFYILIGSLPLLVAL                              TTTQNPTLLNMLTTT 180
[illegible]
Db    402 LITAIY GLEMPLM                  TQRNKLPLHIITTNPIHIR  ELLIMILH 441
Qy    594 LLPRIMAAPQDSL 606
Db    442 MLPLIMLTMKPAL 454
```





```
DR   EMBL; AB042809; BAA95804.1; .
DR   EMBL; AB042524; BAA95649.1; .
DR   InterPro; IPR000847; HTH LysR.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR000260; Oxidored_q5_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF00662; oxidored_q5_N; 1.
DR   PROSITE; PS00044; HTH_LYSR_FAMILY, UNKNOWN 1.
KW   Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
SQ   SEQUENCE   459 AA;  51851 MW;  [illegible] CRC64;

  Query Match            [illegible]%;  Score 133.5;  DB 8;  Length 459;
  Best Local Similarity  [illegible]%;  Pred. No. [illegible];
  Matches  113;  Conservative   82;  Mismatches  174;  Indels  219;  Gaps   27;

Qy     59 LSILVLLLLAMLVRPPLWPCVPCPPCLPCPVLLLACLRPEAVAAVPHVLLSSLCLL 117
Db      6 LPSLMLLPLIWLSSPKKIWIN                    VLSLSFLISLLSLFLLW 43

Qy    118 LPCLALPPLLLASAPSQ[illegible]KIEAPPGAWKLLGLFYYAALYY[illegible]AACAIAGH-------- 169
Db     44 QIDENYKNFTNMF[illegible]     [illegible]      [illegible]IAWLLLLMLMAQ[illegible]NHLKKDNNVL 91

Qy    170       -        IAAHLLGSILSWAHLGVQVWQRAECPQVPKI-    ---------  -YK 202
Db     92 QKLYISMLISL[illegible]LLIMIFSAIEL IMPYILPEAILIPTLIIIIRWGNQTEKLNAGIYFL 150

Qy    203 YYSLLASLPLLIGL---       GFLSLWYPVQLVRSFSKRIGAGSKGLQSSYSEEYLRNLL 255
Db    151 FYILLGSIPLLIALILIQNHVGILNL---- MILSFIIHT--     -LDASWS-     NNLL 195

Qy    256 CRKKLGSSYHISKHGFLSWARVCLRHCIYIPQPGFHLPL-KLVLSATLIGTAIYQVALLL 314
Db    196 ------     ----  -----WLACMMAFLIKMPLYCVHLWLPKAHVEAPIAGSMIL AAILL 236

Qy    315 LVGVVPLIQKVRAGVTTL   V-YLLACFCIVLSEPEQEVVELVKHHLWALLVCYISALVL 372
Db    237 KLG-   SYGMIRISIILDPLIKYMAYPF-ILLS--           LWG           -- 268

Qy    373 [illegible]LLIPLVLMPCLVIHEINLPALHPGAALLLSPLHRGHIPSKQAIPWMSPSAY QTAFI 431
Db    269          - MIMISSI[illegible]LRQIDLKSL  ----          IAYSSVSHMALV 297

Qy    432 [illegible]GLLVLQIIPPL[illegible] -          ---IIALAFLVL-- MPVLHGRNLLLPRSLESSWPPWLI 478
Db    298 IASIMIQIPWSEMGAIMLMIAHGLISSLLP[illegible]LANSNYERIHGRIMIMARGLQMVFPLMAI 357

Qy    479 LALAV[illegible]LNMAAEWV[illegible]LLIHL[illegible]HLLLNRPVLTAA          [illegible]NVL[illegible]GAM[illegible]A 529
Db    358 WWLMACLANLAL            [illegible]NLMGELPTIMGLLSWGN[illegible]LMGTNL[illegible]GMYS 407

Qy    530 IWPVLLS    ALYNAI              HLSQMLLSLLIPRAAILDFGYYI 565
Db    408 MYMILLIQRGKLINHMINLPGHIRELILMALHMIPLLLLLISLKLIT 455


Search completed: August 31, 2002, 14:5[illegible]:27
Job time: 258 sec
```